

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

**Defective images within this document are accurate representations of
the original documents submitted by the applicant.**

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

#33
AP
4/29/02
RECEIVED
TECH CENTER 1600/2900
MAR 15 PM 3:51

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Hein et al.
Title: TRANSGENIC PLANTS
EXPRESSING ASSEMBLED
SECRETORY ANTIBODIES
Appl. No.: 09/200,657
Filing Date: 11/25/98
Examiner: Bui, Phuong T.
Art Unit: 1638

RECEIVED
MAR 20 2002
TECH CENTER 1600/2900

DECLARATION OF RICHARD LERNER, M.D., UNDER 37 C.F.R. § 1.132

I, Richard Lerner, hereby declare that:

1. I was educated in the biomedical sciences at Northwestern University and at Stanford University where at the latter, I received a B.S. and an M.D. degree. I interned at the Palo Alto Stanford Hospital and was a postdoctoral fellow in experimental pathology at the Scripps Clinic and Research Foundation. For the next thirty years, I have held various scientific positions at the now Scripps Research Institute, the world's largest, private non-profit biomedical research facility. During this time, I also held administrative leadership positions at Scripps, the most prominent being chairman of the Department of Molecular Biology from 1982-1986 and then president of the Scripps Research Institute. I have conducted research for over thirty years and am the author or co-author of over 370 published scientific articles in biomedical science. Many of the publications from the late 1980s and early 1990s deal with issues of antibody specificity and antibody expression. I have been a member of the editorial board of more than 10 scientific journals and have been an official reviewer for hundreds articles submitted for publication. A brief summary of my accomplishments and a recent copy of my Curriculum Vitae is attached as APPENDIX 1.

THIS PAGE BLANK (USPTO)

2. I am generally familiar with the work disclosed in the above-captioned patent application and am aware that claims from this application have been rejected by the Examiner as allegedly being obvious over a dissertation by Klaus During ("the During dissertation") of the University of Koln, FRG, carrying the date of July 1988. I am also familiar with the During dissertation (attached as APPENDIX 2),¹ which is poorly organized and quite difficult to read. My previous review of this document has been extended both in terms of scope and depth in order to assist the Examiner in evaluating During's alleged achievements from the perspective of the ordinary skilled artisan at the time period between the alleged publication date of During (i.e., July 1988) and what I have been informed is the earliest filing date of the above referenced patent application (i.e., October 27, 1989). Thus, before discussing the During dissertation, I will review the state of the art of immunoglobulin expression prior to October 27, 1989. As my analysis will demonstrate, there existed at the time of During a significant prejudice in the art against the possibility that plant cells were capable of processing and assembling an immunoglobulin with antigen specificity (i.e., an antibody molecule).

I. There was a significant prejudice in the art against the possibility that plant cells could successfully process and assemble an immunoglobulin with antigen binding specificity around the time of the During dissertation (i.e., circa 1988/1989).

3. It was commonly known prior to October 27, 1989 that antibodies were naturally the product of specialized mammalian cells known as lymphocytes. It was also known that only a subset of lymphocytes, the B cells, produced antibodies. Natural antibodies comprise a four chain structure made up of two identical heavy chains of about 50,000 daltons each and two identical light chains of about 25,000 daltons each. It was appreciated by the early 1980s that the biology of antibody expression was complex and varied with the maturation state of the B cell. For example, it was known that the heavy

¹ The English translation which I reviewed bears the name "Ralph McElroy Translation Company, 910 West Avenue, Austin Texas, Job No. 1596-81522." All reference to pages in During are to this English translation.

THIS PAGE BLANK (USPTO)

and light chain genes were inherited as non-functional gene segments and that only functional heavy and light chain genes were formed by rearrangement in B cells. It was further appreciated that rearrangement of the heavy chain gene preceded that of the light chain gene and that in the earliest member of the B cell lineage, the pre-B cell, only the heavy chain gene was rearranged. Thus, in the pre-B cells, the heavy chain was produced intracellularly and was not transported to the cell surface or secreted. In young B cells where light chain has been productively rearranged, the light chain was synthesized along with the heavy chain and a full-sized immunoglobulin is formed and expressed on the cell surface. When young B cells encounter foreign antigen and receive the proper signals, the B cells mature into a plasma cell, which secretes or exports antibody to the extracellular medium.

4. This story of antibody expression in B cells is further complicated by the involvement of an additional protein called binding immunoglobulin protein or "BiP", discovered in the early 1980s. BiP is located in the endoplasmic reticulum of B cells and binds to heavy chain produced in pre-B cells. In young B cells, heavy chain also binds to BiP, but in the presence of light chain, the heavy chain is released from BiP and the two chains form a fully assembled antibody which is then exported to the cell membrane.² It was also known that unlike the case in pre-B cells, heavy chain production in the absence of light chain production was often fatal in mature B lymphocytes, a phenomenon known as heavy chain toxicity.³ Heavy chain toxicity was recognized by the absence of myeloma cell mutants that produced heavy but not light chains. Prior to October 1989, it was not clear why heavy chains were toxic to B cells.

5. Thus, it should be apparent from this review that by the later 1980s, immunoglobulin expression in B cells was regarded as a highly developed phenomenon which involved a complex interplay between heavy and light chain and other cellular proteins (e.g., BiP) which orchestrate the processing and export leading to secretion of immunoglobulin at various B cell maturation stages. Consequently, it was

² See e.g., Bole et al., J. Cell Biology, 102:1558-1566, 1986, APPENDIX 3; Hass et al., Nature 306:387-389, 1983, APPENDIX 4.

³ See e.g., Hass et al., Proc. Natl. Acad. Sci. USA, 81:7185-7188, 1984, APPENDIX 5, (p.7185, left column).

THIS PAGE BLANK (USPTO)

commonly believed during the early to mid 1980s that immunoglobulin production was a highly specialized function that could only occur in B cells of mammals. As one author wrote:⁴

Because immunoglobulin production is a specialized function of cells of the B-lymphocyte lineage, it is expected that the conditions for proper Ig gene expression will be provided only in appropriate immunocompetant cells.

Such views were commonly held and reflected a prevailing strong prejudice through much of the 1980s against the possibility of using non-B cells to express antibodies.

6. As the commercial value of monoclonal antibodies became evident in the late 1970s and early 1980s, emphasis was placed on producing these antibodies through recombinant expression so that improved features could be engineered into the antibody molecule. Naturally, much effort was directed to achieving recombinant expression of antibodies in transformed B cells. However, a few individuals attempted to express an antibody in microorganisms that were previously used for expressing foreign proteins. Initial attempts to achieve heavy-light chain assembly in *E. coli* bacteria in the early 1980s were unsuccessful.⁵ In 1988, Better et al. obtained heavy-light chain assembly in the periplasmic space of *E. coli* by placing DNA encoding a bacterial signal sequence 5' to the mature sequence of the light and heavy chain encoding DNA.⁶ In 1985, Wood et al. demonstrated synthesis of a heavy and light chain, cleavage of the signal sequence and assembly into an antibody in extracts of the yeast, *Saccharomyces cerevisiae*.⁷ It was noted, however, that the yeast did not secrete active antibody into the culture medium and the majority of the antibody was in an insoluble fraction.⁸

⁴ Ochi et al, Proc. Natl. Acad. Sci. USA, 80:6351-6355, 1984, APPENDIX 6.

⁵ See Boss et al., Nucl. Acids Res. 12:3791-3806, 1984, APPENDIX 7; Cabilly et al., Proc. Natl. Acad. Sci. USA, 81:3273-3277, 1984, APPENDIX 8.

⁶ Better et al., Science 240:1041-1043, 1988), APPENDIX 9.

⁷ Wood et al., Science 240:1041-1043, 1988, APPENDIX 10.

⁸ *Id.* at p.448, right column.

THIS PAGE BLANK (USPTO)

7. Despite the fact that some reports demonstrated antibody assembly in bacteria and yeast, it is my opinion that this would not have reduced the prevailing prejudice in the art that existed up to October 27, 1989 with respect to antibody assembly by plants. This is because plant cells were understood to be distinct from both mammalian cells and single cell microorganisms. For example, many differences between animals and plant cells were known, as summarized in the in the following commentary:⁹

However, in addition to their photosynthetic capacities, plants show important differences from animals that extend to the cellular level: For example, that vascular system of higher plants carries water, minerals absorbed by the roots, and carbohydrates being transported from sites of synthesis in green tissues to sites of use or storage. A variety of small molecules are also translocated including hormones. Unlike in higher animals, the extraceulluar fluids of the vascular system are not rich in macrmolecules, nor is there an abundant population of circulating cells. Paralleling this, plant cell are surrounded by walls, into and out of the cells but severely restrict access of larger molecules to the cell surface and also limit the interactions of adjacent cells with one another. As with bacteria, the walls provide mechanical support so that plants can grow in fresh water, and multicellular forms can circulate a relatively dilute solution; the cells can resist osmotically induced pressures without bursting. These characteristics, plus the presence in the plant of vacuoles controlling ion distribution within the cell, produce balances of ion and water movements that can be quite different from the ones usually found in animal cells.

Although some procaryotes and fungi also have cell walls, plant cell walls are more complex and variable, providing additional functions not present in cell walls of other organisms. Furthermore, processes underlying signal sequence processing and protein secretion also were understood to be different between mammalian cells and plant cells.¹⁰

⁹ See Holtzman et al., *Cell & Organelles*, 3rd ed. 1984, CBS College Publishing, p.357-360, APPENDIX 11.

¹⁰ Duong et al., *J. Biol. Chem.* 262:6328-6333, APPENDIX 12 , 6328-6329 ("Furthermore, the biological properties of one of the synthetic peptide analogs reveals differences in the nature of the mammalian *versus* the plant secretory apparatus.").

THIS PAGE BLANK (USPTO)

In addition, at the time of the During dissertation, it had not been demonstrated that plant cells contained a BiP protein or an analogue that would function equivalently.¹¹ My analysis, therefore, convinces me that a person of skill in the field of immunology or protein expression would have been prejudiced in the time frame up through October 27, 1989 against the possibility that plants could be used to assemble antigen binding immunoglobulins, despite the fact that success was observed in a few selected examples of unicellular microorganisms.

8. It should be apparent from the above that there was a sound basis for a real prejudice in the art against using plants to produced a processed and assembled immunoglobulin which is antigen specific around the time of the During dissertation (*circa* 1988/1989). Were this not the case, then Applicant's invention clearly would not have been roundly hailed in both the scientific literature and in the general press as a significant scientific discovery and medical breakthrough.¹² Because of this prejudice, the ordinary skilled artisan in 1988/1989 would have shown a good deal of skepticism when evaluating evidence supporting a claim of successful antibody assembly in plant cells and would not be convinced unless the proof was well founded. As will be discussed below, the proof supporting During's claims of successful immunoglobulin expression in plants is far below what was needed to overcome the prejudice in the art and reasonably convince one of ordinary skill that plants could be used to process and assemble an immunoglobulin with antigen binding specificity. In fact, During's antibody expression results are so contradictory and uncontrolled that they would not have been convincing even if there were no prejudice in the art.

II. The strategy for light chain and heavy chain expression adopted by During is different from that disclosed by the patent application in question.

9. The During dissertation describes an attempt to co-express the heavy and light chains of an IgM antibody known at B1-8 in plants. His strategy was to

¹¹ See Miernyk et al., J. Cell Biol., 197:abstract No. 4333, 1988, APPENDIX 13 (titled "Is there a BiP-like protein in the endoplasmic reticulum of plant cells?").

¹² See Cover page for Nature, volume 242(6245) and article published in the Los Angeles Times (San Diego County), November 2, 1989 (both attached as APPENDIX 14).

THIS PAGE BLANK (USPTO)

encode a leader sequence in front of the codons for the mature antibody chains. For the heavy chain leader, he introduced nucleic acid encoding the barley alpha amylase signal sequence directly in front of (5' to) the DNA encoding the amino terminal end of the heavy chain.¹³ During used the same amylase signal sequence for the light chain, but this time he included nucleic acid encoding three additional amino acids (Gly-Ser-Met) between the DNA encoding the leader sequence and the DNA encoding the amino terminus of the light chain (*Id.*). During, therefore, chose two distinct strategies for expressing each immunoglobulin chain.

10. The additional amino acids that would, through During's strategy, be encoded at the 3' end of the B1-8 light chain leader sequence were unusual in the context of known eukaryotic signal cleavage sites. At the time of During's dissertation, it was not clear what effect additional amino acids at the end of a leader sequence would have on final processing of the leader. Studies by the present Applicants as well as others in the art indicate that by introducing the amino acids Gly-Ser-Met between the C-terminal end of the leader and the first amino acid of the mature kappa chain, the structure of the potential cleavage site is altered.¹⁴

11. It is now clear from the art that mutations introduced in the vicinity of a cleavage site can adversely influence signal processing. This conclusion is based on analysis of many scientific reports, which address: 1) naturally occurring cleavage sites,¹⁵ and 2) the effect of mutation on the function of a cleavage site.¹⁶ During's strategy to include two polar amino acids (Gly-Ser) followed by a methionine residue to the cleavage site for the light chain leader likely obscured substrate recognition by disrupting structural determinants required for processing activity.¹⁷

¹³ During dissertation, APPENDIX 2, p18, top of page.

¹⁴ See Nielsen et al., Protein Engineering, 10:1-6, 1997, APPENDIX 15 (see abstract).

¹⁵ See references in APPENDIX 15-23 (see abstracts).

¹⁶ See references in APPENDIX 24-38 (see abstracts).

¹⁷ See Nielsen et al., Protein Engineering, 10:1-6, 1997, APPENDIX 15 (see abstract); Duffaud et al., J. Biol. Chem., 263:10224-10228, 1988, APPENDIX 39 (see abstract).

THIS PAGE BLANK (USPTO)

12. In addition, I conclude that During's strategy for mating the light chain leader to the mature light chain sequence cleavage site was likely subject to cleavage site ambiguity. This is reflected in statistical¹⁸ and neural network¹⁹ predictions of probable cleavage sites. For example, in all instances where Gly-Ser-Met is theoretically introduced distal to a cleavage site, an incorrect or absent cleavage site is predicted.²⁰ Whereas these computations are useful tools, they do not guarantee experimental success.²¹ As will be detailed further in my declaration, not surprisingly, the leader sequence strategy used by During did not demonstrate a clear experimental success. In view of the clear deficiencies of During's work, his claims of successful expression would not have been reasonably believed by the ordinary skilled artisan.

III. During's experimental evidence allegedly supporting heavy and light chain processing and assembly of an antigen-specific immunoglobulin is inconsistent and lacks in critical controls.

13. The During dissertation describes that plant cells were initially transfected with DNA encoding the antibody light chain but not the heavy chain. Incredibly, During was unable to detect light chain in the cells.²² The failure to detect expression of light chains would, in my opinion, have been disturbing to the ordinary skilled artisan because it was known that light chains can be readily expressed without heavy chains in B cells. In addition, a very low level of light chain expression would make it that much more difficult to detect heavy-light chain assembly. Furthermore, an increased relative heavy chain expression, which under the circumstances might be necessary to achieve assembly in view of the low levels of expressed light chain, conceivably could result in toxicity if plant cells turned out to be susceptible to heavy

¹⁸ See Von Heijne et al., APPENDIX 23 (see abstract).

¹⁹ See Nielsen et al., APPENDIX 15 (see abstract).

²⁰ See Nielsen et al., APPENDIX 15; Center for Biological Sequence Analysis, Dept. of Biotechnology, The Technical University of Denmark website (www.cbs.dtu.dk/; SignalP cleavage site predictor).

²¹ See Von Heijne et al., APPENDIX 23 (see abstract).

²² During dissertation, APPENDIX 2, p. 80, line 2 ("Repeated attempts to directly detect the light chain of B1-8 and for T4 lysozyme from the crude extract of tobacco mesophyll protoplasts were unsuccessful.").

THIS PAGE BLANK (USPTO)

chain toxicity as is the case for B cells.²³ Thus, During's failure to detect light chain expression in cells transformed to express only the light chain would have raised serious complicating issues requiring a more thorough investigation.

14. In spite of the failure to detect light chain expressed by itself, During proceeded to clone the heavy chain with the light chain into a dual cassette expression vector and attempted to express both chains in plant cells. Page 86-95 of the During dissertation contains the results and discussion related to detection of antibody expression which employed Western blotting and tissue printing (see section 3.14, pages 86 to 90) as well as ultrastructural analysis of transfected plant tissue (see section 3.16, pages 92-95). As indicated by the statement below, During understood that his expression system was suboptimal, and that if assembled antibody were produced, the amount would be very low relative to the total protein.²⁴

Due to the expected level of low percentage of sought protein in the total protein, sensitive detection methods had to be developed. For this purpose, a non-radioactive detection system, which is compatible with the system for DNA hybridization described in 2.7.5 and 3.3, was worked out for detection of proteins on nitrocellulose and Immobilon filters. . . .

A method therefore had to be developed that permits the sought protein to be enriched from the crude extract before Western blot or preferably to be isolated and concentrated to detectable concentration.

The acknowledged deficiencies of During's expression system ultimately forced him to develop a pre-enrichment scheme in order to apply Western blotting. Such scheme would have raised concern with the ordinary skilled artisan as to the overall credibility of During's findings because direct Western blotting was known to be extremely sensitive and had previously been used for demonstrating foreign host expression. Indeed, in contrast to During's inability to use direct Western blotting, assembled antibody generated in tobacco plants using Applicant's system as described in the above-captioned

²³ See Boyle et al, APPENDIX 3.

²⁴ During dissertation, APPENDIX 2, p.86, 2nd to last paragraph and 4th paragraph of p. 87.

THIS PAGE BLANK (USPTO)

patent application were readily detected by direct Western blotting (see Figure 5 of the above-captioned patent application). During's efforts to increase sensitivity creates a greater potential for artifactual results, requiring additional experimental controls. As will be evident below, During's experimental approach wholly fails to provide the additional controls.

15. To enrich for antibody in the plant extract prior to Western blotting, During exposed large volumes of plant extract to several rounds of affinity purification with CNBr activated Sepharose 4b to which is attached Ls136 antibody (monoclonal antibody allegedly specific for the light chain) or NP hapten.²⁵ Bound antibody (if present) was eluted, according to During, with 0.1 M glycine, pH 2.3 presumably for the Ls136-sepharose or NIP-cap for the NP-Sepharose.²⁶ During's Western blotting results are shown in Figure III/232 and discussed on page 89-90. During indicates that direct Western detection was unsatisfactory and that he could detect only "processed light chain" in the callus material through the pre-purified extracts.²⁷ During's statement appears to indicate that heavy chain detection was attempted, but was unsuccessful. I base this observation on the fact that During planned to detect heavy chain with an anti μ antibody by Western blotting²⁸ and because there were no results featured showing heavy chain detection. Thus, During fails to detect either the light or heavy chain in direct Western blotting but allegedly detects a "processed" light chain but not a heavy chain in the indirect (pre-processed) Western blots. To conclude as he does from the Western results that assembled B1-8 antibody was present in the plant extract, During must infer that which he is attempting to prove, that fully assembled antibody must have been present in the extract for light chain to have been enriched following binding to the NP hapten immunoabsorbent.²⁹ As will be seen below, this faulty circular

²⁵ During dissertation, APPENDIX 2, p. 87, 4th full paragraph through page 89, lines 11-13.

²⁶ During dissertation, APPENDIX 2, p.47, lines 7-10.

²⁷ During dissertation APPENDIX 2, p89, lines 3-4

²⁸ During dissertation APPENDIX 2, p.86, lines 12-13.

THIS PAGE BLANK (USPTO)

reasoning is open to alternative explanations which directly conflict with During's conclusion.

16. During's assertion of successful heavy-light assembly severely lacks the type of proof that normally would normally be required under the circumstances to draw such conclusion. Even accepting for the sake of argument During's detection of a "processed light chain" by indirect Western blotting, assembly still requires a heavy chain and During has no direct proof from either the direct or indirect Western that heavy chain was expressed. In fact, any of a number of artifacts may be the cause of During's Western results. For example, assuming for the sake of argument that a light chain was expressed by the plants, the NP hapten immunoabsorbents might have bound light chain even without heavy chain expression during the "enrichment" procedure. This is a real possibility because light chains alone have been known to exhibit a low but real affinity for antigen, particularly in circumstances such as the present case when the antigen is immobilized. Alternatively, light chains could have been purified by the immunoabsorbents if an endogenous plant protein existed which coincidentally had features in common with the heavy chain of the B1-8 antibody. Although I have no evidence that either of these scenarios were at play, they should (and could) have been eliminated by employing any of a number of possible controls. For example, During could have directly demonstrated that heavy chain was absolutely required for light chain binding during the pre-enrichment step. Alternatively, or in addition, During could have used biosynthetic radiolabeling of plant cells in combination with Western blotting to prove that a heavy chain was in fact co-enriched with light chain. This method is well known in the art and was previously used to demonstrate foreign protein expression.³⁰ Biosynthetic radiolabeling also helps to control for stripping of antibody during a low pH elution of an antibody immunoabsorbent column (i.e., the Ls136 adsorbent), a problem

²⁹ During dissertation APPENDIX 2, p.89, lines 14-17 ("By indirect detection of isolation of the functional antibody with hapten [and] detection of the light chain contained [there]in it could be demonstrated that plants are capable of synthesis of active monoclonal antibody whose individual chains are fused with a plant signal peptide.).

³⁰ See e.g., Rothstein et al., *Nature*, 308:662-665, 1984, APPENDIX 40, Fig. 2 (expression of wheat α -amylase in yeast cells labeled with ³⁵S-methionine).

THIS PAGE BLANK (USPTO)

encountered with CNBr.³¹ Since During employed low pH elution and CNBr linkage, he should have provided controls to address this potential problem.³²

17. During also evaluated antibody expression using “tissue printing,” in which a leaf is pressed against a membrane in order to bind various proteins in the leaf to the membrane and the membrane is probed by the various immunological reagents used in Western blotting.³³ No extract is made in this case and no enrichment is used. In these experiments, During allegedly detected light chain, heavy chain and “aggregated B1-8,” (the latter conclusion presumably is a result of Ac38 antibody binding to the tissue print). It is disturbing that During never even attempted to explain why heavy chain was detected in these experiments but not in the Western blots. In any event, even if he could explain this discrepancy in a scientifically acceptable manner, the tissue printing experiment inexplicably lacks the type of controls that are normally required to support the conclusions that During attempts to draw. The controls needed for tissue printing are the same types of controls typically employed in antibody binding assays such as ELISA or immunohistochemistry. As is always the case for antibody reagents (polyclonal or monoclonal), one must use antigen inhibition to validate the specificity of binding in each assay format (i.e. that binding is “antigen-specific”). For example, Gubler et al. demonstrated α -amylase secretion by plant cells by inhibiting tissue staining with antigen.³⁴ Similarly, Wood et al. used antigen inhibition to demonstrate B1-8 antibody assembly in yeast (the same antibody used by During).³⁵ In fact, Wood demonstrated that the NIP form of antigen inhibited better than the NP form of hapten, a well known signature of the B1-8 antibody.³⁶ The During dissertation evidences that During was

³¹ See Lihme et al., J. Chromatography, 376:299-305, 1986, APPENDIX 41, table 1.

³² Ls136 antibody which may have leached into the “enriched” leaf extracts may be detected by cross reaction with the secondary reagents in the western blot.

³³ During dissertation, APPENDIX 2, paragraph bridging p.88 and 89.

³⁴ See Gubler et al., Planta, 168:447-452, 1986, APPENDIX 42, p1448, right hand column (Specificity of antibody staining evaluated by using antibody previously absorbed with a 50 fold excess of antigen).

³⁵ Wood et al., APPENDIX 10, p448, right hand column, middle of page.

³⁶ Wood et al., APPENDIX 10, p448, right hand column, middle of page.

THIS PAGE BLANK (USPTO)

aware of the Wood et al. reference³⁷ and the heteroclitic nature of the B1-8 antibody,³⁸ but for reasons unknown to me, During failed to conduct any antigen inhibition controls. Sossountzov et al. studying immunogold localization of abscisic acid in plant tissues of *chenopodium polyspermum L*, likewise used an antigen inhibition control which the authors recognized as vital (my emphasis added to the following quote).³⁹

Preabsorbed ABA antibodies, **which constitute an essential control for determining the specificity of the immunolabelling**, produced less than one gold particle per μm^2 over cytoplasm, nucleic and plastids.

Although During diluted his antibody reagents with wildtype plant extract and still observed binding, this control is not sufficient to exclude other artifacts. For example, binding directed to insoluble antigens not present in the extract would not be inhibited in the presence of the extract.

18. The During dissertation includes immunogold electron microscopic analysis of the plant cells containing the dual heavy and light chain vector apparently using the same antibody reagents used in the Western blotting and tissue printing experiments. The dissertation describes that the Ac38 antibody (allegedly specific for a properly assembled B1-8 antibody) reacted with endoplasmic reticulum (ER) and chloroplasts in cells from induced plant stems.⁴⁰ Also described is that the Ls136 antibody (allegedly specific for light chains), reacted with cytoplasm, but not with ER or chloroplasts.⁴¹ Further described is the observation that none of the antibody reagents labeled in the vicinity of the plant cell wall or the intercellular space, and that no reaction with the golgi apparatus or vesicles were seen.⁴² During concludes without explanation that these experiments indicate “synthesis and assembling of a monoclonal

³⁷ During dissertation, APPENDIX 2, p.17, first full paragraph.

³⁸ During dissertation, APPENDIX 2, p.86, lines 21-23.

³⁹ Sossountzov et al, Planta, 168:471-481, 1986, APPENDIX 43 (p. 480, left hand column).

⁴⁰ During dissertation, APPENDIX 2, p. 94, first full paragraph (“Each chloroplast contains about 23-30 gold particles, often localized on thylakoid membranes”).

⁴¹ During dissertation, APPENDIX 2, p.94, lines 7-10.

⁴² During dissertation, APPENDIX 2, p.94, lines 10-12.

THIS PAGE BLANK (USPTO)

B1-8 antibody . . . occur [sic] in the plant cells on the rough ER . . . ⁴³ If During's conclusion about successful B1-8 assembly in plant cells was correct, the Ac38 immunoreactivity would mean that assembled antibody, i.e., both a light chain and a heavy chain, were present in the ER and chloroplasts. Such conclusion, however, is inconsistent with During's own failure to observe light chain immunogold labeling of ER or chloroplasts. Furthermore, the chloroplastic reactivity by the Ac38 antibody is perplexing and During provides no explanation for why allegedly assembled antibody would be concentrated in such an organelle⁴⁴ I note that only weak immunogold labeling of chloroplasts was observed in During's plant cells expressing T4 lysozyme.⁴⁵

19. During's alleged success also is in conflict with the failure to detect by immunogold labeling, the Ac38 antibody or light chain reactive antibodies binding in the vicinity of the plant cell wall, the intercellular space, the golgi apparatus or vesicles.⁴⁶ In particular, the presence of processed and assembled immunoglobulin in B cells, or for that matter, protein expressed naturally by plants, was known to involve the golgi and/or vesicles. For example, ultrastructural analysis of immunoglobulin producing B cell lines by Newell et al., showed light chains in the golgi and assembled immunoglobulin in the ER, perinuclear space and/or golgi apparatus.⁴⁷ Others have reported that heavy and light chains are synthesized on separate membrane bound ribosomes and sequestered in the cisternae of the ER from which they are transported into the smooth membranes of the golgi apparatus.⁴⁸ Furthermore, ER and golgi staining was also previously observed for α -amylase in gibberellin-treated barley cells.⁴⁹ In contrast, During observed no golgi immunoreactivity using his light chain, heavy chain or Ac38

⁴³ During dissertation, APPENDIX 2, p.94, last paragraph.

⁴⁴ Chloroplastic targeting was known to require a special amino terminal transport peptide, something which to my knowledge has not been demonstrated in antibodies. Chloroplastic immunoreactivity was observed for abscisic acid but this is not a protein (see Sossountzov et al., APPENDIX 43).

⁴⁵ During dissertation, APPENDIX 2, p.97, line 1.

⁴⁶ During dissertation, APPENDIX 2, p.94, lines 10-12.

⁴⁷ Newell et al., British J. Haematology, 50:45-459, 1982, APPENDIX 44 (see abstract).

⁴⁸ Newell et al., APPENDIX 44 (see p.446).

⁴⁹ Gubler et al., APPENDIX 42, Fig. 2.

THIS PAGE BLANK (USPTO)

antibody reagents. cells were even capable of antibody assembly. Unusual results might be acceptable if plant cells were even capable of antibody assembly in unique and previously unknown ways, however, unusual results cannot make up for the lack of controls in other experiments.

20. The reliance During places on the Ac38 antibody to support his conclusions from the immunogold labeling is misplaced. The Ac38 antibody reacts with the Ac38 idiotype, an antigenic determinant expressed on cell surface immunoglobulin in the B cell population of C57BL/6 mice at very high frequency (around 1/1,000).⁵⁰ It was known, however, in the early 1980s that the majority of Ac38 idiotype bearing antibodies induced in C57BL/6 mice (by immunizing with Ac38 antibody) do not have NP binding specificity.⁵¹ This means that Ac38 antibody binding cannot be used to claim that NP antigen binding specificity is present, even if Ac38 binding were proven to be Ac38 idiotype specific by inhibiting binding with the specific antigen. Thus, even if During had done the proper antigen inhibition controls, he still could not use Ac38 antibody binding to infer that B1-8 light chain and heavy chain were properly processed and assembled resulting in NP antigen specificity. In any event, this issue is moot because During did not even perform the proper controls.

21. The During dissertation concludes that the ultrastructural immunogold results support the Western blotting results which together demonstrate synthesis and assembly of a monoclonal antibody in plants.⁵² Again, no reasoning to support this conclusion is provided and in my opinion, it is wholly unwarranted in view of the experimental difficulties such as the failure to detect the B1-8 heavy chain, and the lack of proper controls.

IV. Conclusion

⁵⁰ Dildrop et al., EMBO, 3:517-523, 1984, APPENDIX 45, p.517, right hand column.

⁵¹ Dildrop et al., APPENDIX 45, p517, right hand column ("The resulting Ac38-positive hybridomas are thus in the majority unreactive with the NP hapten.").

⁵² During dissertation, APPENDIX 2, p.94, last paragraph.

THIS PAGE BLANK (USPTO)

22. My analysis of the immunology and protein expression art, circa 1988/1989, convinces me that a person skilled in such art would not have reasonably believed that plant cells could be used to express a properly processed and assembled antibody molecule. My view is not altered by the limited reports of antibody assembly in single cell microorganisms because such organisms would not have been considered predictive in this context for a plant cell. The art, therefore, supports a prejudice against the possibility of using plant cells to process and assemble a functional antigen-specific immunoglobulin molecule. The During dissertation does not overcome this prejudice because of serious unexplained inconsistencies and the absence of critical controls. In fact, During's work is so deficient that I would make the same conclusion even if there were no prejudice to overcome. Much more was needed to exclude artifactual results, which are more likely to be present when one is working at the limits of detectability. Even in its best light, giving the During dissertation the full benefit of the doubt, I am convinced that the skilled artisan would not have believed During's claims of immunoglobulin assembly in plant cells. Although During was awarded a Ph.D. degree for his dissertation, I do not know what weight his dissertation committee gave to his antibody expression experiments versus his T4 lysozyme expression experiments. I am aware that During eventually published his antibody plant work in a peer-reviewed journal (i.e., *Plant Molecular Biology*) after the inventors of the above-captioned application first published their work.⁵³ In fact, During's publication in *Plant Molecular Biology* discusses the earlier publication by the inventors, Hiatt and Hein as a successful demonstration of antibody expression in plants.⁵⁴ In my opinion, During's antibody experiments would not have been published were they not supported by the earlier published success of inventors, Hiatt and Hein. Had During attempted to publish his work in a peer reviewed journal before Hiatt and Hein published their work, my extensive experience as a reviewer/editor of scientific journals leads me to conclude During's work would most likely have been rejected as inconclusive.

⁵³ During et al., *Plant Mol. Biol.* 15:281-293, 1990, APPENDIX 46.

⁵⁴ During et al., APPENDIX 46, p.291, right column.

THIS PAGE BLANK (USPTO)

23. I am convinced that the ability to process, assemble, and secrete antigen specific immunoglobulin in plants was not achieved prior to the disclosure by Hiatt *et al.*, (see, e.g., article in Nature⁵⁵ and in U.S. Patent No. 5,202,422). The inventors used a different strategy from During and achieved a significant level of expression, allowing detection of the assembled chains by direct Western blotting. Hiatt *et al.*, not During, was the first to convincingly demonstrate the ability of plants to support production of an antigen-specific immunoglobulin in a manner that overcomes the prejudice in the art.

24. It is also my opinion that the During dissertation is silent about the ability of plants to produce single polypeptide forms of an immunoglobulin. Such single polypeptide immunoglobulins generally comprise at least the antigen-binding portion of a heavy chain and the antigen binding portion of a light chain. A commonly known form of single polypeptide immunoglobulin is the single chain Fv (sFv) fragment, which comprises a heavy chain variable region, a light chain variable region, and a short peptide which links the two regions together. I could find nothing in the During dissertation that addresses expression of a single polypeptide form of immunoglobulin such as an sFv fragment. The deficiencies of the During dissertation which has been extensively discussed with respect to a dual chain immunoglobulin apply equally well if not better with a single polypeptide immunoglobulin; The heavy chain variable region in the case of the single polypeptide must still assemble with the light chain variable region portion in order to achieve an antigen specific immunoglobulin. Again it was the inventors of the above-captioned patent application, not During, who were the first describe assembly of a an antigen-specific sFv in plant cells.

25. The During dissertation also fails to teach how to successfully use plant cells to express a heavy chain or light chain polypeptide, but not both, in plant cells. As already discussed, During attempted light chain expression without the heavy chain (but not vice versa) but failed to detect light chains in the plant cells. During did not even attempt to express heavy chains by themselves in plant cells. In contrast, the

⁵⁵ Hiatt *et al.*, Nature, 342:76-78, 1989, APPENDIX 46.

THIS PAGE BLANK (USPTO)

inventors of the above-captioned patent application, were the first describe that plant cells can express the light chain or the heavy chain separately in a plant cell. In my opinion, the ability of plants to express each individual chain (light or heavy) was unexpected, particularly in the case of the heavy chain which was known at least in mature B cells to cause toxicity when expressed without a light chain.

26. I am aware that the Applicants, which I understand to be Andrew Hiatt and Mitch Hein, are inventors of several other pending patent applications in which the claims have been rejected over the During dissertation. I give permission to Applicants and their attorneys to file this declaration in support of any such other applications. In doing so, Applicants and their attorneys may replace the first page of the declaration with a substitute page that is otherwise the same except for the case caption information.

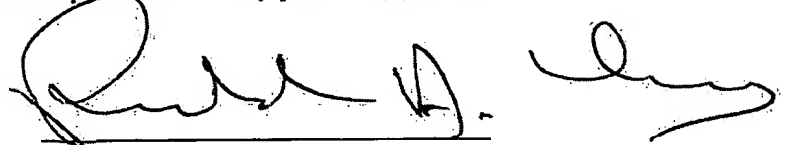
I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true; and further that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001

THIS PAGE BLANK (USPTO)

of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the captioned patent application or any patent issued therefrom.

3/11/02

Date

A handwritten signature in black ink, appearing to read 'Richard Lerner', written over a horizontal line.

Richard Lerner, M.D.
The Scripps Research Institute
10550 North Torrey Pines Rd.
La Jolla, CA 92037

THIS PAGE BLANK (USPTO)

Richard A. Lerner, M.D.

Dr. Richard A. Lerner's 30-year scientific career is particularly significant not only for the broad scope of his achievements in several diverse areas of biomedical research, but for his leadership and vision in concurrently directing the totality of scientific activities at The Scripps Research Institute, the country's largest private, non-profit biomedical research organization. His work spans a wide range of seemingly disparate discoveries, from unique insights into protein and peptide structure to the recent identification of a sleep-inducing lipid. He has been widely recognized by numerous prestigious societies and organizations in the U.S. and abroad.

One of his most recent accomplishments, and that for which he is perhaps most well known, is the groundbreaking work of converting antibodies into enzymes, permitting the catalysis of chemical reactions considered impossible to achieve by classical chemical procedures. While it has taken enzymes acting on natural biological systems millions of years of evolution to reach their present level of efficiency, antibodies can be produced overnight, for obtaining an almost limitless variety of products -- beyond natural ones -- with an efficiency that may exceed that of natural enzymes.

Dr. Lerner's prolific scientific output is accomplished with his simultaneous appointment as President of The Scripps Research Institute. His visionary leadership has kept the Institute at the frontier of science in several explicit and highly focused areas, giving the organization particular strength at the border between chemistry and biology. Emphasizing interdisciplinary collaboration that would be unlikely if not impossible at many major U.S. universities, Dr. Lerner embraces the notion of providing the faculty with a significant degree of freedom and the full range of technical resources to remain at the cutting edge. And recognizing the trend of shrinking Federal resources for science as far back as the early 1980s, Dr. Lerner has encouraged the formation of large-scale industrial collaboration agreements with major pharmaceutical industries, which have given the Institute the opportunity to recruit, build, expand, and remain state-of-the-art in facilities and instrumentation.

Dr. Lerner graduated from Northwestern University and Stanford Medical School. He interned at Palo Alto Stanford Hospital, and received postdoctoral training at Scripps Clinic and Research Foundation in experimental pathology. Since 1970 he has held staff appointments at Wistar Institute in Philadelphia and at the Research Institute of Scripps Clinic (renamed The Scripps Research Institute) in La Jolla. He served as Chairman of the Department of Molecular Biology of the Institute from 1982-1986 prior to assuming the presidency of the organization.

Dr. Lerner has received numerous prizes and awards, including the Parke Davis Award in 1978, John A. Muntz Memorial Prize in 1990, San Marino Prize in 1990, The Burroughs Wellcome Fund and the FASEB, Wellcome Visiting Professor Award in 1990,

THIS PAGE BLANK (USPTO)

The College De France Lectureship in 1991, Arthur C. Cope Scholar Award in 1991, The Tenth Annual Jeanette Piperno Memorial Award in 1991, Sixteenth Annual CIBA-GEIGY Drew Award in Biomedical Research in 1992, Humboldt Research Award in 1994, and the Wolf Prize in Chemistry in 1994-1995, the California Scientist of the Year Award in 1996, and the Coley Award for Distinguished Research in Basic and Tumor Immunology in 1999 .

In addition to Charter Membership in the American Society for Virology, Dr. Lerner holds memberships in the American Society for Experimental Pathology, American Society of Microbiology, New York Academy of Sciences, Biophysical Society, and the Pluto Society. He is on the editorial boards for the Journal of Virology, Molecular Biology and Medicine, Vaccine, In Vivo, Peptide Research, Bioorganic and Medicinal Chemistry Letters, Drug Targeting and Delivery, Senior Contributing Editor to PNAS and Chemistry and Biology, Bioorganic and Medicinal Chemistry, Molecular Medicine, Catalysis Technology and Angewandte Chemie.

Dr. Lerner was elected Foreign Member of the Royal Swedish Academy of Sciences in 1985; Member, National Academy of Science USA in 1991; Member, Scientific Policy Advisory Committee, Uppsala University, Uppsala, Sweden in 1991; Member, Scientific Advisory Board, Economic Development Board, Singapore in 1991; Trustee, The Neurosciences Research Foundation, Inc. in 1992; Member, Advisory Board, Chemical & Engineering News in 1994; Member, ETH Institute of Biotechnology Advisory Board, Zurich, in 1994; Member, Stanford Linear Accelerator Center Scientific Policy Committee, Stanford, CA., 1995-1998; Member, Center for Nanoscale Science and Technology Scientific Advisory Board in 1996; Member, College of Chemistry Advisory Board, University of California, Berkeley, 1996-1997; Member, California Council on Science and Technology Board of Directors, 1996-1997; Member, Advisory Steering Group for Chemistry, California State University, 1996; and Member, Academic Committee of the Board of Governors of Technion Israel Institute of Technology, 1998.

THIS PAGE BLANK (USPTO)

CURRICULUM VITAE
Richard Alan Lerner, M.D.

Date of Birth: August 26, 1938
Place of Birth: Chicago, Illinois
Nationality: American
Marital Status: Married
Home Address: 7750 E. Roseland Drive
La Jolla, CA 92037
Business Address: The Scripps Research Institute
10550 N. Torrey Pines Road
La Jolla, CA 92037
Business Telephone: (858) 784-8265

Education: Northwestern University, 1956-1959
Stanford University Medical School, B.S., M.D. 1959-1964
Internship, Palo Alto Stanford Hospital, 1964-1965

Professional Positions

1965 Research Fellow, Department of Experimental Pathology, Scripps Clinic and Research Foundation, La Jolla, California
1968 Associate, Wistar Institute, Philadelphia, Pennsylvania
1970 Associate, Department of Experimental Pathology, Scripps Clinic and Research Foundation, La Jolla, California
1972 Associate Member, Department of Experimental Pathology, Scripps Clinic and Research Foundation, La Jolla, California
1974 Member, Department of Immunopathology, Scripps Clinic and Research Foundation, La Jolla, California
1977 Member, Department of Cellular and Developmental Immunology, research Institute of Scripps Clinic, La Jolla, California
1980 Head, Committee for the Study of Molecular Genetics and Member, Department of Immunopathology, Research Institute of Scripps Clinic, La Jolla, California
1982 Member, Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California
1982 Chairman, Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California
1988 Professor, Department of Chemistry, Research Institute of Scripps Clinic, La Jolla, California
1987 Director, Research Institute of Scripps Clinic, La Jolla, California
1991 President, The Scripps Research Institute, La Jolla, California
1996 Member, Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California

THIS PAGE BLANK (USPTO)

SOCIETIES AND BOARDS

Member Phi Eta Sigma
Member Alpha Omega Alpha
Member American Society of Nephrology
Member American Association of Immunologists
Member American Society for Experimental Pathology
Member American Society of Microbiology
Member New York Academy of Sciences
Member Biophysical Society
Chairman, International Symposium on the Molecular Basis of Cell-Cell Interaction (1977)
Consultant, Special Virus Cancer Program, National Cancer Institute
Ad hoc Member of Molecular Biology Study Section
Member Pluto Society
Fellowship Screening Committee of the California Division, ACS
Charter Member, American Society for Virology
Member Institute of Medicine Ad hoc Committee of New Research Opportunities in Immunology, National Academy of Sciences
Member Organizing Committee for "Modern Approaches to Vaccines", Cold Spring Harbor (1983)
Member Cancer Preclinical Program Project Review Committee, National Cancer Institute
Member Scientific Policy Advisory Committee, Uppsala University, Uppsala Sweden (1991)
Member Scientific Advisory Board Economic Development Board, Singapore (1991)
Member National Academy of Science, USA (1991)
Trustee, The Neurosciences Research Foundation, Inc. (1992)
Member Advisory Board Chemical and Engineering News, Washington, DC (1994-1996)
Member ETH Institute of Biotechnology Advisory Board, Zurich Switzerland (1994)
Member Stanford Linear Accelerator Center Scientific Policy Committee, Stanford, CA (1995-1998)
Member Center for Nanoscale Science and Technology Scientific Advisory Board, Houston, Texas (1996)
Member College of Chemistry Advisory Board, University of California, Berkeley, Berkeley, CA (1996)
Member California Council on Science and Technology Board of Directors, Newport Beach, CA (1996-1997)
Member, Advisor Steering Group for Chemistry, California State University, Los Angeles, Los Angeles, CA (1996)
Member, Editorial Advisory Board for the American Peptide Society Journal, Tucson, Arizona (1997)
Member, Academic Committee of the Board of Governors of Technion-Israel Institute of Technology, Haifa, Israel (1998)
Member, Editorial Advisory Board for Accounts of Chemical Research Journal, Los Angeles, CA (1998-2001)
Member, International Board of Governors, The Peres Center for Peace, Tel Aviv, Israel (1998)
Chairman, Governor's Council on Bioscience, Governor Gray Davis, Sacramento, CA (1999)

THIS PAGE BLANK (USPTO)

Member, The U.S. Department of Energy Biological and Environmental Research Advisory Committee, Washington, DC (1999).

Member, American Academy of Arts and Sciences, Cambridge, MA (2000)

HONORS and PRIZES

- 1970 NIH AID Career Development Award
- 1978 Parke Davis Award
- 1985 Elected Foreign Member, The Royal Swedish Academy of Sciences
- 1986 Recipient of Lita Annenberg Hazen Professor of Immunochemistry
- 1986 Decorated as oficial de La Orden de San Carlos by the President of Colombia
- 1989 Cecil H. and Ida M. Green Chair in Chemistry
- 1990 Honorary Doctorate of Medicine, Karolinska Institute
- 1990 John A. Muntz Memorial Prize, Albany Medical College, New York
- 1990 San Marino Prize, Republic of San Marino
- 1990 The Burroughs Wellcome Fund and the FASEB Wellcome Visiting Professor Award in Basic Medical Science
- 1991 The College De France Lectureship, College De France, Paris
- 1991 Arthur C. Cope Scholar Award in Chemistry, Washington, D.C.
- 1992 CIBA-Geigy Drew Prize in Chemical Science, Drew University, Madison, New Jersey
- 1993 Hyp J. Dauben Award in Chemistry, University of Washington, Seattle, Washington
- 1994 Humboldt Research Award, Bonn Germany
- 1994/95 The Wolf Prize in Chemistry
- 1996 California Scientist of the Year Award
- 1999 William B Coley Award for Distinguished Research in Basic and Tumor Immunology, Cancer Research Institute, New York, NY
- 1999 Windaus-Medal/Award, Georg-August-Universität Göttingen, Göttingen Germany

LECTURESHIPS

- 1984 Casper Wistar Lecture, The Wistar Institute, Pennsylvania.
- 1986 Fourteenth Peter Anthony Leermakers Symposium, Wesleyan University, Connecticut.
- 1988 The Thirteenth Annual B.R. Baker Memorial Lecture University of California, Santa Barbara, California.
- 1988 Allied Corporation Lecture, Rutgers University, New Jersey.
- 1988 10th Annual Jim Mc Ginnis Memorial Lecture, Duke University, North Carolina.
- 1988 Selected to give the address on the occasion of Michael Heidelberger's 100th Birthday, New York University, N.Y.

THIS PAGE BLANK (USPTO)

- 1988 Clark Lecture, San Jose State University, California.
- 1988 Dow Chemical Matrix Lecture, Midland, Michigan.
- 1988 Abbott Lecture, Brandeis University, Waltham, Massachusetts.
- 1989 Rennebohm Lecture, University of Wisconsin, Wisconsin.
- 1989 Mobay Lecture, University of Pittsburgh, Pittsburgh, PA.
- 1989 Spink Lecture, University of Minnesota, Minnesota.
- 1990 Smith Kline Lecture, Yale University, Connecticut.
- 1990 The John A. Muntz Lecture, Albany Medical College, Albany, N.Y.
- 1990 The First Josef Fried Symposium, University of Chicago, Illinois.
- 1991 Wallace P. Rowe Symposium, National Institute of Health, Bethesda, Maryland.
- 1991 The Mabel Beckman Lectures in Technologies, Stanford University, Stanford, California.
- 1991 Bruce Merrifield's 70th Birthday Symposium, Rockefeller University, New York, NY.
- 1991 Lilly Symposium, New York, NY.
- 1991 Peter Reichard Symposium, Stockholm Sweden.
- 1991 John G. Reinhold Lecture, University of Pennsylvania, Pennsylvania.
- 1991 The Tenth Annual Jeanette Piperno Memorial Award, Temple University, Philadelphia, Pennsylvania.
- 1992 Schering Lectures, Schering, Berlin, West Germany.
- 1992 Biomega Lectures in Organic Chemistry, McGill University, Montreal, Quebec, Canada.
- 1992 Seventh Annual William S. Johnson Symposium in Organic Chemistry, Stanford University, Stanford California.
- 1992 Sixteenth Annual CIBA-GEIGY Drew Award in Biomedical Research, Drew University, Madison, New Jersey.
- 1993 The John Howard Appleton Lecture in Chemistry, Brown University, Providence, Rhode Island.
- 1993 August-Wilhelm-von-Hofmann-Lecture, Gesellschaft Deutscher Chemiker, Germany.
- 1994 Lemieux Lecture, University of Alberta, Edmonton, Canada.
- 1995 Bender Lectureship in Chemistry, Northwestern University, Evanston, Illinois.
- 1995 Alfred Burger Lecturer in Chemistry, University of Virginia, Charlottesville, Virginia.
- 1996 Ephraim Katzir's 80th Birthday Symposium, The Israel Academy of Sciences and Humanities, Jerusalem, Israel
- 1996 Falk-Plaut Lecture, Columbia University, New York, NY.
- 1996 Debye Lectures, Cornell University, New York, NY.
- 1996 Harvey Lecture, The Rockefeller University, New York, NY.
- 1996 Honors Program Lecture, New York University School of Medicine, New York, NY.
- 1997 The Twelfth Biennial Carl S. Marvel Symposium, University of Arizona, Tucson, AZ.
- 1997 John and Dolores Stille Science Symposium, Colorado State University, Fort Collins, Colorado.

THIS PAGE BLANK (USPTO)

- 1997 Roma Eisenstark Memorial Lecture in Cancer Research, University of Missouri-Columbia, Columbia, MO.
- 1997 Arrhenius Lecture in Chemistry, Stockholm University, Stockholm, Sweden.
- 1998 Zeneca Pharmaceuticals Distinguished Lecturer in Chemistry, Wilmington, Delaware.
- 1999 Twenty-Seventh Annual Verna and Marrs McLean Science and Society Lecture in Biochemistry, Baylor College of Medicine, Houston, Texas.
- 1999 21st Oliver H. Lowry Lecture, Washington University School of Medicine, St. Louis, Missouri.
- 1999 Adolf-Windaus-Gedächtnis-Lecture, Georg-August-Universität Göttingen, Göttingen Germany.
- 2000 Rodney Porter Memorial Lecture, University of Oxford, Oxford, UK.
- 2000 Beverly and Raymond Sackler Lecture, Memorial Sloan-Kettering Cancer Center, New York, NY.
- 2000 Sumner Lecture, Cornell University, Ithaca, NY.
- 2001 18th Herbert C. Brown Lecture, Purdue University, West Lafayette, IN.
- 2001 Linus Pauling Lecture at the Centennial Celebration, California Institute of Technology, Pasadena, California.

Editorial Boards:

Journal of Virology
Molecular Biology and Medicine
Vaccine
Journal of Peptide Research
Bioorganic and Medicinal Chemistry Letters
Drug Targeting and Delivery
Chemistry and Biology
Bioorganic and Medicinal Chemistry
Molecular Medicine
Catalysis Technology
Angewandte Chemie

THIS PAGE BLANK (USPTO)

BIBLIOGRAPHY

1. Lerner, R.A. and F.J. Dixon. Transfer of ovine experimental allergic glomerulonephritis (EAG) with serum. *J. Exp. Med.* 124:431-442, 1966.
2. Lerner, R.A. and F.J. Dixon. Induced and spontaneous glomerulonephritis in sheep. *Fed. Proc.* 25:660, 1966.
3. Lerner, R.A. and F.J. Dixon. Spontaneous glomerulonephritis in sheep. *Lab. Invest.* 15:1279-1288, 1966.
4. Lerner, R.A. and F.J. Dixon. Experimental and human glomerulonephritis associated with antiglomerular basement membrane antibodies. *Proc. Natl. Acad. Sci. USA* 15:94, 1967.
5. Lerner, R.A., J.J. Glassock, and F.J. Dixon. The role of antiglomerular basement membrane antibody in the pathogenesis of human glomerulonephritis. *J. Exp. Med.* 126:989, 1967.
6. Unanue, E.R., R.A. Lerner, and F.J. Dixon. Experimental allergic glomerulonephritis (EAG) induced by antibodies to renal basement membrane. *Proc. Vth International Immunopathology Symposium*, P. Grabar and P. Miescher, eds. Schwabe and Co., Basel, 1967.
7. Lerner, R.A., F.J. Dixon, F.E. Young, and R.J. Glassock. Isolation and characterization of soluble human glomerular basement membrane (HuGBM) antigens and the demonstration of pathogenic anti-glomerular antibody in man. *Fed. Proc.* 26:743, 1967.
8. Lerner, R.A. and F.J. Dixon. The induction of acute glomerulonephritis in rabbits by injection of autologous and homologous glomerular basement membrane (GBM) antigens isolated from normal urine. *Proc. First Ann. Conf. of the Am. Soc. of Nephrology*, September, 1967.
9. Glassock, R.J., R.A. Lerner, and F.J. Dixon. Transfer of glomerulonephritis from man to monkey. *Proc. First Ann. Conf. of the Am. Soc. of Nephrology*, September, 1967.
10. Lerner, R.A. and F.J. Dixon. Renal antigens in urine. In: *Proteins in normal and pathological urine*. S. Karger, Basel, 1968.
11. Lerner, R.A., J.J. McPhaul, Jr. and F.J. Dixon. Soluble glomerular basement membrane (GBM) antigens in normal urine. Possible autoimmunogens in man. *Annals of Int. Med.* 68:249, 1968.
12. Starzl, T.E., R.A. Lerner, F.J. Dixon, C.G. Groth, L. Brettschneider, and P.I. Terasaki. The Schwartzman reaction after human renal homotransplantation. *New. Eng. J. Med.* 278:642, 1968.

THIS PAGE BLANK (USPTO)

13. Lerner, R.A. and F.J. Dixon. The induction of acute glomerulonephritis in rabbits with soluble antigens isolated from normal homologous and autologous urine. *J. Immunol.* 100:1277, 1968.
14. Lerner, R.A., F.J. Dixon and S. Lee. Spontaneous glomerulonephritis in sheep. II. Studies on natural history, occurrence in other species and pathogenesis. *Am. J. Path.* 53:501, 1968.
15. Lerner, R.A. Pathogenesis and incidence of spontaneous glomerulonephritis (SG) in sheep and other ruminants. *Fed. Proc.* 27:363, 1968.
16. McPhaul, J.J. and R.A. Lerner. Glomerular basement membrane antigen excretion in normal and nephritic men. *Fed. Proc.* 27:545, 1968.
17. Dixon, F.J., J.J. McPhaul, and R.A. Lerner. Recurrence of glomerulonephritis in the transplanted kidney. *Arch. Int. Med.* 123:554, 1969.
18. Dixon, F.J., J.J. McPhaul, and R.A. Lerner. The contribution of kidney transplantation to the study of glomerulonephritis. The recurrence of glomerulonephritis in renal transplants. *Transp. Proc.* 1:194, No. 1, 1969.
19. Rolley, R.T., G.M. Williams, R.A. Lerner, G. Hanscom, and D.M. Hume. Characterization of antibodies following human homograft rejection. *Transp. Proc.* 1:275, No. 1, 1969.
20. Lerner, R.A. and L.D. Hodge. Nonpermissive infections of mammalian cells: Synthesis of influenza virus in genome in HeLa cells. *Proc. Natl. Acad. Sci. USA* 64:544, 1969.
21. Lerner, R.A., J.J. McPhaul, and F.J. Dixon. Glomerulonephritis mediated by anti-glomerular basement membrane antibodies. International Congress of Nephrology, Stockholm, Sweden.
22. Lerner, R.A. and L.D. Hodge. Gene expression in synchronized lymphocytes: Studies on the control of synthesis of immunoglobulin polypeptides. *J. Cell Physiol.* 77:265, 1971.
23. Lerner, R.A., P.J. McConhey, and F.J. Dixon. Quantitative aspects of plasma membrane associated immunoglobulin in clones of diploid human lymphocytes. *Science* 173:60, 1971.
24. Lerner, R.A., W. Meinke, and D.A. Goldstein. Membrane associated DNA in the cytoplasm of diploid human lymphocytes. *Proc. Natl. Acad. Sci. USA* 68:1212, 1971.
25. Lerner, R.A., W. Meinke, and D.A. Goldstein. Evidence for DNA associated with cytoplasmic membranes of diploid human lymphocytes. *Biophysical J.* 11:48a, 1971.

THIS PAGE BLANK (USPTO)

26. McConahey, P.J., R.A. Lerner, and F.J. Dixon. Quantitation of plasma membrane and cytoplasmic Ig of human diploid lymphocyte clones. *Fed. Proc.* 30:588, 1971.
27. Lerner, R.A., M.B.A. Oldstone, and N.R. Cooper. Cell cycle dependent immune lysis of Moloney virus transformed lymphocytes: Presence of viral antigen, accessibility to antibody and complement activation. *Proc. Natl. Acad. Sci. USA* 68:2584, 1971.
28. Hall, M.R., W. Meinke, D.A. Goldstein, and R.A. Lerner. Synthesis of cytoplasmic membrane associated DNA in lymphocyte nucleus. *Nature New Biology* 234:277, 1971.
29. Tubergen, D.G., J.D. Feldman, E.M. Pollock, and R.A. Lerner. Production of macrophage migration inhibition factor by continuous cell lines. *J. Exp. Med.* 135:255, 1972.
30. Lerner, R.A. Relationship of events at the lymphocyte cell surface to gene expression: Approaches to the problem. *Cont. Topics in Immunochem.* Vol. 1, F.P. Inman, ed., Plenum Press, N.Y., pp. 111-143, 1972.
31. Lerner, R.A., P.J. McConahey, I. Jansen and F.J. Dixon. Synthesis of plasma membrane associated and secretory immunoglobulin in diploid lymphocytes. *J. Exp. Med.* 135:136, 1972.
32. Tan, E.M. and R.A. Lerner. An immunologic Study of the fates of nuclear and nucleolar macromolecules during the cell cycle. *J. Mol. Biol.* 68:107, 1972.
33. Wiktor, T.J., R.A. Lerner, and H. Koprowski. Inhibitory effect of passive antibody on active immunity induced against rabies by vaccination. *Bull. Wld. Hlth. Org.* 45:747, 1972.
34. Ferrone, S., B. Del Villano, M.A. Pellgrio, R.A. Lerner, and R.A. Reisfeld. Expression of HL-A antigens on the surface of cultured human lymphoid cells: Effects of inhibitors of protein and nucleic acid synthesis. *Tissue Antigens* 2:447, 1972.
35. Lerner, R.A., F. Jensen, S. Kennel, F.J. Dixon, G. Des Roches, and U. Francke. Karyotypic, virologic and immunologic analyses of two continuous lymphocyte lines established from New Zealand black mice: Possible relationship of lymphocyte mosaicism to autoimmunity. *Proc. Natl. Acad. Sci. USA* 69:2965, 1972.
36. Del Villano, B.C. and R.A. Lerner. Further studies on plasma membrane associated nucleic acids in cultured human lymphocytes. *Fed. Proc.* 31, 1972.
37. Hall, M.R., D.A. Goldstein, and R.A. Lerner. Effects of inhibitors of macromolecular synthesis on cytoplasmic membrane-associated DNA. *Amer. Soc. Microbiol. Phila.*, 1972.

THIS PAGE BLANK (USPTO)

38. Goldstein, D.A., W. Meinke, M.R. Hall, and R.A. Lerner. Characterization of cytoplasmic membrane-associated DNA in eucaryotic cells. Int. Biophysical Congress, Moscow, 1972.
39. Lerner, R.A. Plasma membrane-associated DNA in diploid human lymphocytes. Symp. on Genes and Antibodies, Scand. J. of Immunol. June, 1972.
40. Lerner, R.A. In: Cell surface alteration as a result of malignant transformation: II. MSS Information Corporation, New York, pp. 52-62, 1972.
41. Lerner, R.A. and F.J. Dixon. The human lymphocyte as an experimental animal. Scientific Am. 228:82, 1973.
42. Kennel, S.J. and R.A. Lerner. Isolation and characterization of plasma membrane associated immunoglobulin from cultured human diploid lymphocytes. J. Mol. Biol. 76:485, 1973.
43. Meinke, W., M.R. Hall, D.A. Goldstein, D.E. Kohné, and R.A. Lerner. Physical properties of cytoplasmic membrane-associated DNA. J. Mol. Biol. 78:43, 1973.
44. Kennel, S.J., B.C. Del Villano, R.L. Levy, and R.A. Lerner. Properties of an oncornavirus glycoprotein: Evidence for its presence on the surface of virions and infected cells. Virology, 55:464, 1973.
45. Kennel, S.J., B.C. Del Villano, and R.A. Lerner. Approaches to the quantitation and isolation of immunoglobulin associated with plasma membranes. In: Methods in Molecular Biology, Vol. 6, T. Zachria, ed. Marcel Dekker, Inc., pp. 1-38, 1973.
46. Hampar, B., K. Rand, R. Lerner, B. Del Villano, R. McAllister, L. Martos, J. Derge, C. Long, and R. Gilden. Formation syncytia in human lymphoblastoid cells infected with wile-type C viruses. Virology 55:453, 1973.
47. Lerner, R.A., F.J. Dixon, B.P. Croker, B.C. Del Villano, F.C. Jensen, S.J. Kennel, and P.J. McConahey. The possible role of oncomaviruses in the etiology and pathogenesis of murine lupus. Adv. in the Biosciences, Pergamon Press, pp. 356-365, 1974.
48. Kingsbury, D.T. and R.A. Lerner. Encapsulation of lymphocyte DNA by vesicular stomatitis virus. Proc. Natl. Acad. Sci. USA 71:753, 1974.
49. Hall, M.R., W. Meinke, D.A. Goldstein, and R.A. Lerner. Differential effects of metabolic inhibitors on cytoplasmic membrane associated and nuclear DNA. Biochem. Biophys. Res. Comm. 60:96, 1974.

THIS PAGE BLANK (USPTO)

50. Leavitt, R.W., R.B. Sparks, D.R. Helsinki, and R.A. Lerner. Immunoelectron microscopic localization of single-stranded regions in DNA. Biochemistry/Biophysics Meeting, 1974.
51. Dixon, F.J., B.P. Croker, B.C. Del Villano, F.C. Jensen, and R.A. Lerner. The possible role of oncornaviruses in murine lupus. Proc. Int. Symp. on Infection and Immunology in the Rheumatic Diseases, Blackwell Scientific, 1974.
52. Croker, B.P., B.C. Del Villano, F.C. Jensen, and R.A. Lerner. Immunopathogenicity and oncogenicity of murine leukemia viruses. I. Induction of immunologic disease and lymphoma in (BALB/c x NZ B) F1 mice by Scripps leukemia virus. J. Exp. Med. 140:1028-1047, 1974.
53. Puga, A., A.B. Jensen, J. Boaz, F. Jensen, D. Kohne, and R.A. Lerner. Molecular analysis of a murine leukemia virus produced by continuously growing thymocytes. Progress in Immunology II, Vol 5, 1974.
54. Dixon, F.J., B.P. Croker, B.C. Del Villano, F.C. Jensen, and R.A. Lerner. Immunopathologic consequences of murine oncornavirus infection. R.F. Beers, Jr. and E.G. Bassett, eds., Raven Press, New York, 1976.
55. Dixon, F.J., B.P. Croker, B.C. Del Villano, F.C. Jensen, and R.A. Lerner. Oncornavirus infection and "auto" immune complex disease of mice. Proc. 2nd Int. Cong. of Immunology, Brighton, England, 1974.
56. Dixon, F.J., B.P. Croker, B.C. Del Villano, R.A. Lerner, and H. Marquardt. Oncornaviruses and immunologic disease. Vth Lepetit Colloquium, Madrid, Spain, 1974.
57. Lerner, R.A. Regulatory macromolecules of the plasma membrane: X. Interphase between thymic differentiation markers and oncornavirion polypeptides. Vth Lepetit Colloquium, Madrid, Spain, 1974.
58. Del Villano, B.C., B. Nave, B.P. Croker, R.A. Lerner, and F.J. Dixon. The oncornavirus glycoprotein gp69/71: A constituent of the surface of normal and malignant thymocytes. J. Exp. Med. 141:172-187, 1975.
59. Siu, C-H, R.A. Lerner, R.A. Firtel, and W.F. Loomis, Jr. Developmentally related expression of a plasma-membrane associated carbohydrate-binding protein in Dictyostelium discoideum. Biophysical J. 15(2):219a, 1975.
60. Kennel, S.J., Del Villano, B.C., Levy, R.L., Lerner, R.A. Biochemical and immunologic characterization of an oncornavirus glycoprotein. Univ. of Toyko Press. pp. 329-338, 1975.

THIS PAGE BLANK (USPTO)

61. Levy, R.L., R.A. Lerner, and F.J. Dixon. The neonatal state is not obligatory for infection of mice with murine leukemia virus. *Fed. Proc.* 34(3):974, 1975.
62. Lerner, R.A. and B.C. Del Villano. Oncornaviral genes and murine differentiation markers. VIIth International Symposium on Comparative Research on Leukemia and Related Diseases, Copenhagen, Denmark, 1975.
63. Del Villano, B.C., B.P. Croker, P.J. McConahey, F.C. Jensen, R.A. Lerner, and F.J. Dixon. Immunopathogenicity and oncogenicity of murine leukemia viruses. II. Expression of muLV p30 and induction of lymphatic leukemia in mice and rats. *Am. Journal of Path.* 82(2):299-313, 1976.
64. Siu, C-H, R.A. Lerner, R.A. Firel, and W.F. Loomis. Changes in plasma membrane proteins during development of Dictyostelium discoideum. In: *Developmental Biology*, Vol. 2, D. McMahon and F. Fox, eds., Benjamin Co., 1976, pp.129-134.
65. Levy, R.L., M.H. Barrington, R.A. Lerner, and F.J. Dixon. On the mechanism of infectivity of a murine leukemia virus in adult mice. In: *Biology of Radiation Carcinogenesis*, J.M. Yuhas, R.W. Tennant, and J.D. Regan, eds., Raven Press, New York, pp. 275-286, 1976.
66. Levy, R.L., R.A. Lerner, and F.J. Dixon. Enhancement of infectivity and oncogenicity of a murine leukemia virus in adult mice by X-irradiation. *Cancer Research* 36:2090-2095, 1976.
67. Siu, C-HY, R.A. Lerner, G. Ma, R.A. Firtel, and W.F. Loomis. Developmentally regulated proteins of the plasma membrane of Dictyostelium discoideum. I. The carbohydrate binding protein. *J. Mol. Biol.* 100:157-178, 1976.
68. Lerner, R.A., C.B. Wilson, B.C. Del Villano, P.J. McConahey, and F.J. Dixon. Endogenous oncornaviral gene expression in adult and fetal mice: Quantitative, histologic, and physiologic studies of the major viral glycoprotein, gp70. *J. Exp. Med.* 143:151-166, 1976.
69. Del Villano, B.C. and R.A. Lerner. Relationship between the oncornavirus gene product gp70 and a major protein secretion of the mouse genital tract. *Nature* 259:497-499, 1976.
70. Del Villano, B.C., S.J. Kennel, and R.A. Lerner. Biological and structural pleomorphism of the oncornavirus envelope glycoprotein, gp70. In: *Contemporary topics in Immunobiology*, Vol. 6, M.G. Hanna and F. Rapp, eds., Plenum Press, New York, 1976.
71. Lerner, R.A., B.C. Del Villano, and S.J. Kennel. Viral genes and cell surface differentiation proteins of the mouse. *Proceedings, ICN-UCLA Winter Conference on Molecular and Cellular Biology, Keystone, Colorado. J. Supramol. Struc.* 4, 1976.

THIS PAGE BLANK (USPTO)

72. Geltosky, J.E., C-H Siu, and R.A. Lerner. Glycoproteins of the plasma membrane of Dictyostelium discoideum during development. *Cell* 8:391-396, 1976.
73. Del Villano, B.C., R.A. Lerner, S.J. Kennel, and J.H. Elder. Expression of an oncornavirus gene product in the mouse genital tract. In: *Proceedings of the Third Annual NICHD Workshop on the Testis*, University of Pittsburgh School of Medicine, eds., Raven Press, New York, 1976.
74. Siu, C-H, R.A. Lerner, and W.F. Loomis. Rapid accumulation and disappearance of plasma membrane proteins during development of wild-type and mutant strains of Dictyostelium discoideum. *J. Mol. Biol.* 116:469-488, 1977.
75. Elder, J.H., F.C. Jensen, M.L. Bryant, and R.A. Lerner. Polymorphism of the major envelope glycoprotein (gp70) of murine C-type viruses: Virion associated and differentiation antigens encoded by a multi-gene family. *Nature* 267:23-28, 1977.
76. Geltosky, J.E., J. Ray, and R.A. Lerner. Use of common plant lectins for isolation and characterization of constitutive and developmentally regulated cell surface associated glycoproteins of Dictyostelium discoideum. In: *Cell Surface Carbohydrates and Biological Recognition*, V. Marchesi, V. Ginsburg, P. Robbins, C. Fox, eds., Alan R. Liss, Inc., New York, pp. 613-619, 1977.
77. Elder, J.H., R.A. Pickett, J. Hampton, and R.A. Lerner. Radioiodination of proteins in single polyacrylamide gel slices: Tryptic peptide analysis of all the major members of complex multicomponent systems using microgram quantities of total protein. *J. Biol. Chem.* 252:6510- 6515, 1977.
78. Bryant, M.L., B.K. Pal, M.B. Gardner, J.H. Elder, F.C. Jensen, and R.A. Lerner. Structural analysis of the major envelope glycoprotein (gp70) of the amphotropic and ecotropic type C viruses of wild mice. *Virology* 84:348-358, 1978.
79. Siu, C-H, R.A. Lerner, and W.F. Loomis. Plasma membrane proteins in wild type and cascade arrested mutant strains of Dictyostelium discoideum. In: *The Molecular Basis of Cell-Cell Interaction*, R.A. Lerner and D. Bergsma, eds., Alan R. Liss, Inc. New York, 1978, pp. 439-458.
80. Siu, C-H, R.A. Lerner, and J.E. Geltosky. Altered cell-cell recognition in a temperature conditional mutant strains of Dictyostelium discoideum. In: *The Molecular Basis of Cell-Cell Interaction*, R.A. Lerner and D. Bergsma, eds., Alan R. Liss, Inc. New York, 1978, pp. 459-472.
81. Elder, J.H., F.C. Jensen, J.W. Gautsch, R.A. Lerner, and M. Vogt. Structural analysis of the surface glycoproteins (gp70's) of recombinant murine C-type viruses: Evidence for envelope gene recombination. In: *Advances in Comparative Leukemia Research*, Bentvelzen et al., eds., N. Holland Biomedical Press, 1977.

THIS PAGE BLANK (USPTO)

82. Bordier, C., W.F. Loomis, J. Elder, and R.A. Lerner. The major developmentally regulated protein complex in members of Dictyostelium. J. Biol. Chem. 253:5133-5139, 1978.
83. Elder, J.H., J.W. Gautsch, F.C. Jensen, R.A. Lerner, J.W. Hartley, and W.P. Rowe. Biochemical evidence that MCF murine leukemia viruses are envelope (env) gene recombinants. Proc. Natl. Acad. Sci. USA 74:4676-4680, 1977.
84. Elder, J.H., J.W. Gautsch, F.C. Jensen, and R.A. Lerner. Multigene family of endogenous retroviruses: Recombinant origin of diversity. Invited editorial, J. of the Nat. Cancer Inst. 61:625-638, 1978.
85. Elder, J.H., J.W. Gautsch, F.C. Jensen, and R.A. Lerner. Generation of diversity among murine C-type viruses via envelope gene recombination. In: Differentiation of Normal and Neoplastic Hematopoietic Cells, Cold Spring Harbor Laboratory, 1977, pp. 553-560.
86. Levy, R.L., M.H. Barrington, R.A. Lerner, G.F. Griffin, and C.E. Whitmire. Immunosuppressive effects of 3-methyl-cholanthrene given intratracheally in various strains of mice. Cancer Res. 37:3892-3894, 1977.
87. Gautsch, J.W., J.H. Elder, J. Schindler, F.C. Jensen, and R.A. Lerner. Structural markers on MuLV p30: Functional correlation with Fv-1 tropism. Proc. Natl. Acad. Sci. USA 75:4170-4174, 1978.
88. Gautsch, J.W., R.A. Lerner, D. Howard, Y. Teramoto, and J. Schlom. Strain specific markers for the major structural proteins of highly oncogenic murine mammary tumor viruses by tryptic peptide analyses. J. of Virol. 27:688-699, 1978.
89. Takemoto, L.J., C.F. Fox, F.C. Jensen, J.H. Elder, and R.A. Lerner. Nearest-neighbor interactions of RNA tumor virus glycoprotein, gp70, on the surface of murine cells. Proc. Natl. Acad. Sci. 75:3644-3648, 1978.
90. Schlom, J., W. Drohan, Y. Teramoto, P. Hand, D. Colcher, R. Callahan, G. Todaro, K. Kufe, D. Howard, J. Gautsch, R. A. Lerner, and G. Schidlovsky. Diversity of mouse mammary tumor virus genetic information and gene products in rodents. In: Workshop on the Origins of Inbred Mice, H.C. Morse III, ed., Academic Press, New York, 1978.
91. Oskarrson, M., J.H. Elder, J.W. Gautsch, R.A. Lerner, and G.F. Vande Woude. Chemical determination of the m1Moloney sarcoma virus pP60^{gag} gene order: Evidence for unique peptides in the C-terminus of the polyprotein. Proc. Natl. Acad. Sci. USA 75:4694-4698, 1978.

THIS PAGE BLANK (USPTO)

92. Fischinger, P.J., A.E. Frankel, J.H. Elder, R.A. Lerner, and J.N. Ihle. Biological, immunological, and bio-chemical evidence that HIX virus is a recombinant between Moloney leukemia virus and a murine xenotropic C type virus. *Virology* 90:241-254, 1978.
93. Lerner, R.A. Recombinant origins of leukemogenic murine viruses. Parke, Davis Lecture. *The Am. J. of Path.* 93:10-20, 1978.
94. Ray, J., T. Shinnick, and R.A. Lerner. A mutation altering the function of a carbohydrate binding protein blocks cell-cell cohesion in developing Dictyostelium discoideum. *Nature* 279:215-221, 1978.
95. Elder, J.H., J.W. Gautsch, F.C. Jensen, R.A. Lerner, T.M. Chused, H.C. Morse, J.W. Hartley, and W.P. Rowe. Differential expression of two distinct xenotropic viruses in NZB mice. *Clinical Immunology and Immunopathology* 15:493-501, 1980.
96. Geltosky, J.E., J. Weseman, A. Bakke, and R.A. Lerner. Identification of a cell surface glycoprotein involved in cell aggregation in Dictyostelium discoideum. *Cell* 18:391-398, 1979.
97. Gautsch, J., J. Elder, F. Jensen, and R. Lerner. In vitro construction of B-tropic virus by recombination: B-tropism is a cryptic phenotype of xenotropic murine retroviruses. *Proc. Natl. Acad. Sci. USA* 77:2989-2993, 1980.
98. Sutcliffe, J.G., T.M. Shinnick, R.A. Lerner, P. Johnson, and I. Verma. Nucleotide sequence of a cloned murine leukemia viral DNA. *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. 44, 1980.
99. Shinnick, T.M. and R.A. Lerner. The cbpA gene: Role of the 26,000 dalton carbohydrate binding protein in intracellular cohesion of developing Dictyostelium discoideum cells. *Proc. Natl. Acad. Sci. USA* 77:4788-4792, 1979.
100. Sutcliffe, J.G., T.M. Shinnick, I.M. Verma, and R.A. Lerner. Nucleotide sequence of Moloney leukemia virus: The 3' end reveals details of replication, analogy to bacterial transposons, and an unexpected gene. *Proc. Natl. Acad. Sci. USA* 77:3302-3306, 1980.
101. Geltosky, J.E., C.R. Birdwell, J. Weseman, and R.A. Lerner. A glycoprotein involved in aggregation of Dictyostelium discoideum is distributed on the cell surface in a non-random fashion favoring cell junctions. *Cell* 21:339-345, 1980.
102. Gautsch, J.W., J.H. Elder, F.C. Jensen, and R.A. Lerner. Structural diversity among retroviral gene products: A molecular approach to the study of biological function through structural variability. *Adv. in Cancer Res.* 34:171-209, 1980.

THIS PAGE BLANK (USPTO)

103. Sutcliffe, J.G., T.M. Shinnick, N. Green, F-T Liu, H.L. Niman, and R.A. Lerner. Chemical synthesis of a protein predicted from nucleotide sequence allows detection of a new retroviral gene product. *Nature* 287:801-805, 1980.
104. Sutcliffe, J.G., T.M. Shinnick, and R.A. Lerner. Moloney leukemia virus is a transposon: Nucleotide sequence analysis identifies genes and replication details. *Cold Spring Harbor Symp. on Quant. Biol.* Vol. XLV, 1980, pp. 707-710.
105. Bakke, A. and R.A. Lerner. The cascade of membrane events during development. *Subcellular Biochem.* 8:75-122, 1981.
106. Lerner, R.A., J.G. Sutcliffe, and T.M. Shinnick. Antibodies to chemically synthesized peptides predicted from DNA sequences as probes of gene expression. Mini-review, *Cell* 23:309-310, 1981.
107. Ray, J. and R.A. Lerner. A biologically active receptor for the carbohydrate binding protein(s) of Dictyostelium discoideum. *Cell* 28:91-98, 1982.
108. Lerner, R.A., N. Green, H. Alexander, F-T Liu, J.G. Sutcliffe, and T.M. Shinnick. Chemically synthesized peptides predicted from the nucleotide sequence of the hepatitis B virus genome elicit antibodies reactive with the native envelope protein of Dane particles. *Proc. Natl. Acad. Sci. USA* 78:3403-3407, 1981.
109. Green, N., T.M. Shinnick, O. Witte, A. Ponticelli, J.G. Sutcliffe, and R.A. Lerner. Sequence-specific antibodies reveal that the maturation of Moloney leukemia virus envelope polyprotein involves the removal of a carboxy-terminal peptide. *Proc. Natl. Acad. Sci. USA* 78:6023-6027, 1981.
110. Ray, J., T.M. Shinnick, S. Alexander, and R.A. Lerner. Genetics and biochemistry of cell recognition molecules in Dictyostelium discoideum. *J. of Supramolec. Struct.* Alan R. Liss, Inc. N.Y., pp. 789-793, 1981.
111. Schindler, J., J. W. Gautsch, R.A. Lerner, and N. Hopkins. Biochemical analysis of the p30s of N-, B-, and B-NB-tropic murine leukemia viruses of BALB/c origin. *J. of Virol.* 39:703-712, 1981.
112. Shinnick, T.M., R.A. Lerner, and J.G. Sutcliffe. Nucleotide sequence of Moloney murine leukemia virus. *Nature* 293:543-548, 1981.
113. Lerner, R.A., N. Green, A. Olson, T. Shinnick, and J.G. Sutcliffe. The development of synthetic vaccines. *Hospital Practice* 16:55-62, 1981.
114. Lerner, R.A. Antibodies to chemically synthesized peptides from DNA sequences as probes of gene expression. In: *Molecular Genetic Neuroscience*, F. E. Bloom and S. Bird, eds., Raven Press, New York, 1982.

THIS PAGE BLANK (USPTO)

115. Green, N., H. Alexander, A. Olson, S. Alexander, T.M. Shinnick, J.G. Sutcliffe, and R.A. Lerner. The immunogenic structure of the influenza virus hemagglutinin. *Cell* 28:477-489, 1982.
116. Bittle, J.L., R.A. Houghten, H. Alexander, T.M. Shinnick, J.G. Sutcliffe, R.A. Lerner, D.J. Rowlands, and F. Brown. Protection against foot-and-mouth disease by immunization with a chemically synthesized peptide predicted from the viral nucleotide sequence. *Nature* 298:30-33, 1982.
117. Gerin, J.L., H. Alexander, J.W-K Shih, R.H. Purcell, G. Dapolito, R. Engle, N. Green, J.G. Sutcliffe, T.M. Shinnick, and R.A. Lerner. Chemically synthesized peptides of hepatitis B surface antigen duplicate d/y specificities and induce subtype-specific antibodies in chimpanzees. *Proc. Natl. Acad. Sci. USA* 80:2365-2369, 1982.
118. Sutcliffe, J.G., R.J. Milner, R.E. Bloom, and R.A. Lerner. Common 82 nucleotide sequence unique to brain RNA. *Proc. Natl. Acad. Sci. USA* 79:4942-4946, 1982.
119. Levy, D.E., R.A. Lerner, and M.C. Wilson. A genetic locus regulates the expression of tissue specific mRNAs from multiple transcription units. *Proc. Natl. Acad. Sci. USA* 79:5823-5827, 1982.
120. Lerner, R.A. Tapping the immunological repertoire at will to produce antibodies of predetermined specificity. *Nature* 299:592-596, 1982.
121. Alexander, S., H. Alexander, N. Green, and R.A. Lerner. Antisera against synthetic peptides neutralize influenza viruses of differing hemagglutinin subtypes. *Proc. Natl. Acad. Sci. USA*, 1982.
122. Sutcliffe, J.G., T.M. Shinnick, N. Green, and R.A. Lerner. Antibodies that react with predetermined sites on proteins. *Science* 219:6660-666, 1983.
123. Lerner, R.A. Synthetic vaccines. *Scientific American* 248(2):66-74, 1983.
124. Luka, J., L. Sternas, H. Jornvall, G. Klein, and R.A. Lerner. Antibodies of predetermined specificity of the N-terminus of a cellular "p53" protein react with the native molecule: Evidence for the presence of different "p53s." *Proc. Natl. Acad. Sci. USA* 80:1199-1203, 1982.
125. Shinnick, T.M., J.G. Sutcliffe, N. Green, and R.A. Lerner. Synthetic peptide immunogens as vaccines. *Ann. Rev. Micro.* 37:425-446, 1983.
126. Ray, J. and R.A. Lerner. Cellular interaction in the morphogenesis of Dictyostelium discoideum. In: *Cell Receptors and Cell Communication in Invertebrates*, A.H. Greenberg, ed., Marcel Dekker, Inc., New York, 1983.

THIS PAGE BLANK (USPTO)

127. Alexander, S., T.M. Shinnick, and R.A. Lerner. Mutants of Dictyostelium discoideum blocked in expression of all members of the developmentally regulated discoidin multi-gene family. *Cell* 34:467-475, 1983.
128. Niman, H., R. Houghten, L. Walker, R. Reisfeld, I. Wilson, J. Hogle, and R. Lerner. Generation of reactive antibodies by short peptides is an event of high frequency: Implications for the structural basis of immune recognition. *Proc. Natl. Acad. Sci. USA* 80:4949-4953, 1983.
129. Alexander, H. and R.A. Lerner. Chemically synthesized peptide analogs of the hepatitis B surface antigen. In: *Advances in Hepatitis Research*, ed. F. Chisari, Masson Pub. Inc., 1983, pp. 223-229.
130. Seiden, M.V., B. Clevinger, T. Srouji, J.M. Davie, S. McMillan, and R.A. Lerner. A synthetic peptide induces a new anti-dextran idotype. *Annals of Immunology*, 135C:77-82, 1983.
131. Alexander, S., A.M. Cibulsky, and R.A. Lerner. Ion dependence of the discoidin I lectin from Dictyostelium discoideum. *Differentiation* 24:209-212, 1983.
132. Clarke, B.E., A.R. Carroll, D.J. Rowlands, B.H. Nicholson, R.A. Houghten, R.A. Lerner, and F. Brown. Synthetic peptides mimic subtype specificity of foot-and-mouth Disease virus. *Federation of European Biochemical Societies*. Vol. 157, No. 2, 1983.
133. Alexander, H., D. Johnson, N. Green, R. Lerner. Mimicking alloantigens with chemically synthesized peptides differing in single amino acids. *Nature* 306:697-699, 1983.
134. Rowlands, D., B. Clarke, A. Carroll, F. Brown, B. Nicholson, J. Bittle, R. Houghten, and R. Lerner. The chemical basis for variation in the major antigenic site eliciting neutralizing antibodies in foot-and-mouth disease virus. *Nature* 306:694-697, 1983.
135. Milich, D.R., G.G. Leroux-Roels, D.L. Peterson, R.A. Lerner, and F.V. Chisari. Dissection of the immunogenicity of the hepatitis B surface antigen into T and B cell recognition sites. *Nature*, 1983.
136. Lerner, R.A. Antibodies of Predetermined Specificity in Biology and Medicine. *Proceedings of the XVIIIth Solvay Conference on Chemistry*. George Van Binst, ed., 43-49, (Springer-Verlag, Berlin) 1973.
137. Shinnick, T.M., J.G. Sutcliffe, J.L. Gerin, R.H. Purcell, J.L. Bittle, H. Alexander, D.J. Rowlands, F. Brown, and R.A. Lerner. The chemical nature of antiviral serotype specificity. In: *International Symposium of Affinity Chromatography and Biological Recognition*, I. Chaiken, ed., Academic Press, 1984, pp. 343-353.

THIS PAGE BLANK (USPTO)

138. MacFarlan, R.I., B. Dietzschold, H. Koprowski, J.G. Sutcliffe, and R.A. Lerner. Localization of the immunodominant domains of rabies virus glycoprotein. In: *Modern Approaches to Vaccines*, R.M. Chanock and R.A. Lerner, eds., 1984, p. 139.
139. Chen, P.P., R.A. Houghten, S. Fong, G.H. Rhodes, T.A. Gilbertson, J.H. Vaughan, R.A. Lerner, and D.A. Carson. Anti-hypervariable region antibody induced by a defined peptide. A new approach for studying the structural correlates of idiotypes. *Proc. Natl. Acad. Sci. USA* 81:1784, 1983.
140. McMillan, S., M.V. Seiden, R.A. Houghten, B. Clevinger, J.M. Davie, and R.A. Lerner. Synthetic idiotypes: The third hypervariable region of murine anti-dextran antibodies. *Cell* 35:859-863, 1983.
141. Milner, R.J., F.E. Bloom, C. Lai, R.A. Lerner and J.G. Sutcliffe. Brain specific genes have ID sequences in their introns. *Proc. Natl. Acad. Sci. USA* 81:713, 1983.
142. Shinnick, T.M., J.G. Sutcliffe, and R.A. Lerner. Antibodies to synthetic immunogens as probes for virus protein expression and function. In: *Concepts in Viral Pathogenesis*, A. Notkins and M.B.A. Oldstone, eds., 1984, pp. 361-365.
143. Shinnick, T.M., J.G. Sutcliffe, N. Green, and R.A. Lerner. Peptide elicited protein-reactive antibodies in molecular biology and medicine. *J. Invest. Derm.* 83:112s-115s, 1983.
144. Bittle, J.L., P.A. Worrell, R.A. Houghten, R.A. Lerner, D.J. Rowlands, and F. Brown. Immunization against foot-and-mouth disease with a chemically synthesized peptide. In: *Modern Approaches to Vaccines*, R.M. Chanock and R.A. Lerner, eds., Cold Spring Harbor Press, 1984, p. 93.
145. Rowlands, D.J., B.E. Clarke, A.R. Carroll, F. Brown, B.H. Nicholson, J.L. Bittle, R.A. Houghten, and R.A. Lerner. Chemical basis for variation in the major antigenic site eliciting neutralizing antibodies in foot-and-mouth disease virus. In: *Modern Approaches to Vaccines*, R.M. Chanock and R.A. Lerner, eds., Cold Spring Harbor Press, 1984, p. 87.
146. Shinnick, T.M., J.G. Sutcliffe, N. Green, J.L. Bittle, H. Alexander, R.A. Lerner, J.L. Gerin, D.J. Rolands, and F. Brown. Molecular basis of antiviral serotype specificity. In: *Modern Approaches to Vaccines*, R.M. Chanock and R.A. Lerner, eds., Cold Spring Harbor, 1984, p. 65.
147. Seiden, M.V., B. Clevinger, S. McMillan, A. Srouji, R. Lerner, and J. Davie. Chemical synthesis of idiotypes: Evidence that antisera to the same JH₁ peptide to detect multiple binding site-associated idiotypes. *J. Exp. Med.* 159:1338-1350, 1984.

THIS PAGE BLANK (USPTO)

148. Lerner, R.A. Antibodies of predetermined specificity. *Advances in Immunology* 36:1-44, 1984.
149. Gerin, J.L., R.A. Lerner, and R.H. Purcell. Alternative sources of hepatitis B vaccine. *Develop. Biol. Standard* 54:135-137, 1983.
150. Wilson, I.A., H.L. Niman, R.A. Houghten, A.R. Cherenon, M.L. Connolly, and R.A. Lerner. The structure of an antigenic determinant in a protein. *Cell* 37:767-778, 1984.
151. Sutcliffe, J.G., R.J. Milner, J.M. Gottesfeld, and R.A. Lerner. Identifier sequences are transcribed specifically in brain. *Nature* 308:237-241, 1984.
152. Dillner, J., L. Sternas, B. Kallin, H. Alexander, Ehlin-Henriksson, H. Jornvall, G. Klein, and R.A. Lerner. Antibodies against a synthetic peptide identify the Epstein-Barr virus determined nuclear antigen. *Proc. Natl. Acad. Sci. USA* 81:4652-4656, 1984.
153. Lamb, J.R., E.D. Zanders, P. Lake, R.G. Webster, D.D. Eckels, J.N. Woody, N. Green, R.A. Lerner and M. Feldmann. Inhibition of T cell proliferation by antibodies to synthetic peptides. *Eur. J. Immunol.* 14:153-157, 1984.
154. Gerin, J.L., R.H. Purcell, and R.A. Lerner. Recombinant DNA and synthetic peptide approaches to HBV vaccine development: Immunogenicity and protective efficacy in chimpanzees. In: *Modern Approaches to Vaccines*, R.A. Chanock and R.A. Lerner, eds., Cold Spring Harbor Press, 1984, p. 115.
155. Niman, H.L., B.R. Hough-Evans, V.D. Vacquier, R.J. Britten, R.A. Lerner, and E.H. Davidson. Proteins of the sea urchin egg vitelline layer. *Developmental Biology* 102:390-401, 1984.
156. Shinnick, T.M. and R.A. Lerner. The structure of proteins in solution as revealed by antibodies to predetermined sites. *Immunology Convocation*, 1984.
157. Tainer, J.A., E.D. Getzoff, H. Alexander, R.A. Houghten, A.J. Olson, R.A. Lerner, and W.A. Hendrickson. The reactivity of anti-peptide antibodies is a function of the atomic mobility of sites in a protein. *Nature* 312:127-134, 1984.
158. Levy, D., R. Lerner, and M. Wilson. Gv-1 locus coordinately regulates the expression of multiple endogenous murine retroviruses. *Cell*, 1984.
159. Shinnick, T.M. and R.A. Lerner. The immunology of synthetic peptides. In: *Contemporary Topics in Molecular Immunology*, 1984.
160. Tainer, J.A., E.D. Getzoff, Y. Paterson, A.J. Olson, and R.A. Lerner. The atomic mobility component of protein antigenicity. *Ann. Rev. Immunol.* 3:501-535, 1985.

THIS PAGE BLANK (USPTO)

161. Getzoff, E.D., J.A. Tainer, and R.A. Lerner. The chemistry of antigen-antibody union. Proceedings of the Sixteenth International Leucocyte Culture Conference, 1984.
162. Lamb, J.R., M. Feldman, N. Green, and R.A. Lerner. The influence of antigen conformation on the activation and induction of unresponsiveness in cloned human T lymphocytes, 1984.
163. Gerin, J.L., R.H. Purcell, and R.A. Lerner. The use of synthetic peptides to identify protective epitopes of the hepatitis B surface antigen. Vaccines, 1984.
164. Moriarty, A.M., H. Alexander, G.B. Thornton, and R.A. Lerner. Antibodies to peptides detect new hepatitis B antigen: Serological correlation with hepatocellular carcinoma. Science 227:429-433, 1985.
165. Lerner, R.A. and T.M. Shinnick. The structure of proteins in solution as revealed by antibodies to predetermined sites. 9th Int. Convoc. Immunol., Buffalo, New York, 1984, p. 50.
166. MacFarlan, R.I., B. Dietzschold, T.J. Wiktor, M. Kiel, R. Houghten, R.A. Lerner, J.G. Sutcliffe and H. Koprowski. T-cell responses to cleaved rabies virus glycoprotein and to synthetic peptides. J. Immunol. 133:2745, 1984.
167. Shinnick, T.M., J.G. Sutcliffe, and R.A. Lerner. The structure and functional determinants of proteins as revealed by antibodies to predetermined sites. In: Trends in Pharmacological Sciences, 1984.
168. MacFarlan, R.I., B. Dietzschold, H. Koprowski, M. Kiel, R. Houghten, R.A. Lerner and J.G. Sutcliffe. Localization of the immunodominant domains of rabies virus glycoprotein. In: Modern Approaches to Vaccines, R.M. Chanock and R.A. Lerner, eds., Cold Spring Harbor Press 1984, pp. 139-143.
169. Levy, D.E., R.A. Lerner, and M.C. Wilson. The Gv-1 locus coordinately regulates the expression of multiple endogenous murine retroviruses. Cell 41:89, 1985.
170. Levy, D.E., R.A. Lerner, and M.C. Wilson. Normal expression of polymorphic endogenous retroviral RNA containing segments identical to MCF virus. J. Virology 56:691-700, 1985.
171. Wilson, I.A., D.H. Haft, E.D. Getzoff, J.A. Tainer, R.A. Lerner, and S. Brenner. Identical short peptide sequences in unrelated proteins can have different conformations: A testing ground for theories of immune recognition. Proc. Natl. Acad. Sci. 82:5255-5259, 1985.

THIS PAGE BLANK (USPTO)

172. Shinnick, T.M., J.G. Sutcliffe, and R.A. Lerner. The structure and functional determinants of proteins as revealed by antibodies to predetermined sites. In: *Modern Biotechnology and Health: Perspectives for the Year 2000*, ed., M. Patarroyo, Academic Press, Inc., Orlando, Florida, 1986.
173. Dillner, J., B. Kallin, G. Klein, H. Jornvall, H. Alexander, and R.A. Lerner. Antibodies against synthetic peptides react with the second Epstein-Barr virus associated nuclear antigen. *The EMBO J.* 4:1813-1818, 1985.
174. Houghten, R.A., R.A. Lerner, S.R. Hoffmann, P.A. Worrell, P. Wright, and F.A. Klipstein. Completely synthetic vaccine effective against heterologous enterotoxigenic *Escherichia coli*. In: *Vaccines 85, Molecular and Chemical Basis of Resistance to Parasitic, Bacterial, and Viral Diseases*, R.A. Lerner, R.M. Chanock, and F. Brown, eds., Cold Spring Harbor, 1985, pp. 91-94.
175. Francis, M.J., C.M. Fry, D.J. Rowlands, F. Brown, J.L. Bittle, R.A. Houghten, and R.A. Lerner. Priming with peptides of foot-and-mouth disease virus. In: *Vaccines 85, Molecular and Chemical Basis of Resistance to Parasitic, Bacterial, and Viral Diseases*, R.A. Lerner, R.M. Chanock, and F. Brown, eds., Cold Spring Harbor, 1985, pp. 203-210.
176. Perry, N.R., E.J. Ouldrige, P.V. Barnett, D.J. Rowlands, F. Brown, J.L. Bittle, R.A. Houghten, and R.A. Lerner. Identification of neutralizing epitopes of foot-and-mouth disease virus. In: *Vaccines 85, Molecular and Chemical Basis of Resistance to Parasitic, Bacterial, and Viral Diseases*, R.A. Lerner, R.M. Chanock, and F. Brown, eds., Cold Spring Harbor, 1985, pp. 211-216.
177. Milich, D.R., D.L. Peterson, G.G. Leroux-Roels, R.A. Lerner, and F.V. Chisari. Genetic regulation of the immune response to hepatitis B surface antigen (HBsAg). *J. Immunol.* 134:4203, 1985.
178. Dyson, H.J., K.J. Cross, R.A. Houghten, I.A. Wilson, P.E. Wright, and R.A. Lerner. The immunodominant site of a synthetic immunogen has a conformational preference in water for a type II reverse turn. *Nature* 318:480-483, 1985.
179. Dillner, J., L. Eliasson, L. Sternas, B. Kallin, G. Klein, and R.A. Lerner. The use of antibodies against synthetic peptides for studying the Epstein-Barr virus nuclear antigen. In: *EBV and Associated Diseases*, P. Levine et al, eds., Martinus Nijhoff Publishing, Boston, 1985, pp. 446-455.
180. Francis, M.J., C.M. Fry, D.J. Rowlands, F. Brown, J.L. Bittle, R.A. Houghten, and R.A. Lerner. Immunological priming with synthetic peptides of foot-and-mouth disease virus. *J. Gen. Virol.* 66:2347-2354, 1985.

THIS PAGE BLANK (USPTO)

181. Dyson, H.J., K.J. Cross, J. Ostresh, R.A. Houghten, I.A. Wilson, P.E. Wright, and R.A. Lerner. Selection by site-directed antibodies of small regions of peptides which are ordered in water. "Synthetic peptides as antigens," Ciba Foundation Symposium 119, John Wiley & Sons, 1986, Chichester, pp. 58-75.
182. Niman, H.L., A.M.H. Thompson, A. Yu, M. Markman, J.J. Willems, K.R. Herwig, N.A. Habib, C.B. Wood, R.A. Houghten, and R.A. Lerner. Anti-Peptide peptide antibodies detect oncogene-related proteins in urine. *Proc. Natl. Acad. Sci. USA* 82:7924-7928, 1985.
183. Getzoff, E.D., J.A. Tainer, and R.A. Lerner. The chemistry of antigen-antibody union. In: *Immune Regulation*, M. Feldmann and N.A. Mitchison, eds., The Humana Press, 1985, pp. 243-258.
184. Wright, P.E., H.J. Dyson, M. Rance, J. Ostresh, R.A. Houghten, I.A. Wilson, and R.A. Lerner. Antipeptide antibodies and the disorder-order phenomenon. In: *New Approaches to Immunization*, F. Brown, R.M. Chanock, and R.A. Lerner, eds., Cold Spring Harbor Press, New York, 1986, pp. 15-19.
185. Seiden, M.V., R. Heuckeroth, B. Clevinger, S. McMillan, R. A. Lerner, and J.M. Davie. Hypervariable region peptides variably induce specific anti-Idiotypic idiotype antibodies: An approach to determining antigenic dominance. *J. Immunol.* 136:582-587, 1986.
186. Dyson, H.J., M. Rance, R.A. Houghten, P.E. Wright, and R.A. Lerner. The order-disorder paradox in antigen-antibody union: Anti-peptide antibodies as a probe for structural regions of small peptides. "Biological Organization: Macromolecular Interactions at High Resolution," Eds. R.M. Burnett and H.J. Vogel, Academic Press, pp. 227-234, 1986.
187. Dillner, J., B. Kallin, H. Alexander, I. Ernberg, M. Uno, Y. Ono, G. Klein, and R. Lerner. An Epstein-Barr virus (EBV) - determined nuclear antigen (EBNA-5) partly encoded by the transformation-associated Bam WYH region of EBV DNA: Preferential expression in lymphoblastoid cell lines. *Proc. Natl. Acad. Sci. USA* 83:6641-6645, 1986.
188. Kohan, S., B. Froya, E. Cederlund, T. Fairwell, R.A. Lerner, J. Johansson, S. Khan, M. Ritzen, H. Jornvall, S. Cekan, and E. Diczfalusy. Peptides of postulated inhibit activity. Lack of in Vitro inhibit activity of a 94-residue peptide isolated from human seminal plasma, and of a synthetic replicate of its C-terminal 28-residue segment. *FEBS* 199:242-248, 1986.
189. Tramontano, A., K.D. Janda, and R.A. Lerner. Chemical reactivity at an antibody binding site elicited by mechanistic design of a synthetic antigen. *Proc. Natl. Acad. Sci. USA* 83:6736-6740, 1986.

THIS PAGE BLANK (USPTO)

190. Ouldrige, E.J., N.R. Parry, P.V. Barnett, C. Bolwell, D.J. Rowlands, F. Brown, J.L. Bittle, R.A. Houghten, and R.A. Lerner. Comparison of the structures of the major antigenic sites of foot-and-mouth disease viruses of two different serotypes. In: *New Approaches to Immunization*, F. Brown, R.M. Chanock, R.A. Lerner, eds., Cold Spring Harbor Pres, New York, 1986, pp. 45-49.
191. Ackermann, M., J. Chou, M. Sarmiento, R.A. Lerner, and B. Roizman. Identification by antibody to a synthetic peptide of a protein specified by a diploid gene located in the terminal repeats of the L component of herpes simplex virus genome. *J. Virology* 58:843-850, 1986.
192. Francis, M.J., C.M. Fry, D.J. Rowlands, F. Brown, J.L. Bittle, R. Houghten, and R.A. Lerner. Immune response to uncoupled peptides of foot-and-mouth disease. *J. Immunology* 61:1-6, 1987.
193. Shinnick, T.M. and R.A. Lerner. Anti-peptide antibodies: Some practical considerations. In: *Concepts in Clinical Pathogenesis II*, Springer-Verlag, 1986, pp. 331-335.
194. Shinnick, T.M. and R.A. Lerner. Site directed antibodies in biology and medicine. In: *A Revolution in Biotechnology*, J.L. Marx, ed., Cambridge University Press, 1986.
195. Sternas, L., L. Eliasson, R.A. Lerner, and G. Klein. Quantitation of Epstein-Barr virus (EBV)-determined nuclear antigen (EBNA) by a two-site enzyme immunoassay, in parallel with EBV-DNA. *J. Immunological Methods* 89:151-158, 1986.
196. Dillner, J., R. Szigeti, W. Henle, G. Henle, R.A. Lerner, and G. Klein. Cellular and humoral immune responses to synthetic peptides deduced from the amino acid sequences of Epstein-Barr virus-encoded proteins in EBV transformed cells. *International J. Cancer*.
197. Tramontano, A., K.D. Janda, and R.A. Lerner. Catalytic antibodies. *Science* 234:1566-1570, 1986.
198. Norrby, E., M.A. Mufson, H. Alexander, R.A. Houghten, and R.A. Lerner. Site directed serology with synthetic peptides representing the large glycoprotein (G) of respiratory syncytial virus. *Proc. Natl. Acad. Sci. USA* 84:6572-6576, 1987.
199. Geysen, H.M., J.A. Tainer, S.J. Rodda, T.J. Mason, H. Alexander, E.D. Getzoff, and R.A. Lerner. Chemistry of antibody binding to a protein. *Science* 235:1184-1190, 1987.
200. Getzoff, E.D., H.M. Geysen, S.J. Rodda, H. Alexander, J.A. Tainer, and R.A. Lerner. Mechanisms of antibody binding to a protein. *Science* 235:1191, 1987.
201. Ray, J. and R.A. Lerner. Cellular interaction in the Morphogenesis of Dictyostelium discoideum. In: *Invertebrate Models: Cell Receptors and Cell Communication*, Greenberg, ed., S. Karger, Basel, Switzerland, 1987, pp. 118-142.

THIS PAGE BLANK (USPTO)

202. Wright, P.E., H.J. Dyson, M. Rance, R.A. Houghten, and R.A. Lerner. Identification of folded structures in immunogenic peptides by 2D NMR spectroscopy. In: *Protides of the Biological Fluids*, vol. 35, 1987, pp 477-480.
203. Lai, E.C.H., E.A. Kabat, J. Meienhofer, E.P. Heimer, A.J. Olson, and R.A. Lerner. Inhibition of phosphorylcholine binding to antibodies using synthetic peptides. *Nature* 325:168, 1987.
204. Arrhenius, T., R.A. Lerner, and A.C. Satterthwait. The chemical synthesis of structured peptides using covalent hydrogen-bond mimics. In: *Protein Structure and Design (UCLA Symposia on Molecular and Cellular Biology)*, Volume 69, D. Oxender (ed.), Alan R. Liss, Inc., New York, NY 1987.
205. Norrby, E., G. Biberfeld, F. Chiodi, A. Von Gegerfeldt, A. Nacler, E. Parks, and R.A. Lerner. Antibodies to HIV and to related retroviruses using site-Directed directed serology. *Nature* 329:248-250, 1987.
206. Lerner, R.A., and A. Tramontano. Antibodies as enzymes. *TIBS* 12:427-430, 1987.
207. Napper, A.D., S.J. Benkovic, A. Tramontano, and R.A. Lerner. A stereospecific cyclization catalyzed by an antibody. *Science* 237:1041-1043, 1987.
208. Tramontano, A. and R.A. Lerner. Catalytic antibodies. *Scientific American* 256:58-70, 1988.
209. Johnson, P.R., D.E. Parks, E. Norrby, R.A. Lerner, R.H. Purcell, and R.M. Chanock. Site-directed ELISA identifies a highly antigenic region of the simian immunodeficiency virus transmembrane glycoprotein. *AIDS Res. and Human Retroviruses* 4:159-164, 1988.
210. Norrby, E., G. Biberfeld, F. Chiodi, M.A. Mufson, E. Parks, and R.A. Lerner. Type-specific site-directed HIV serology. *CSH Vaccines*, 1988.
211. Ebina, S., R.A. Lerner, and P.E. Wright. Chemical modification of bovine pancreatic trypsin inhibitor for single site coupling of immunogenic peptides for NMR conformational analysis. *J. Biol. Chem.* 264:7882-7888, 1989.
212. Dyson, H.J., R.A. Lerner, and P.E. Wright. On the physical basis for induction of protein-reactive anti-peptide antibodies. *Ann. Rev. Biophysics* 17:305- 324, 1988.
213. Tramontano, A., K.D. Janda, A.D. Napper, S.J. Benkovic, and R.A. Lerner. Catalytic Antibodies. *CSH Vaccines Vol. III*, pp. 91-96, 1988.

THIS PAGE BLANK (USPTO)

214. Ricksten, A., B. Kallin, H. Alexander, J. Dillner, R. Fahraevs, G. Klein, R.A. Lerner, and L.Rymo. The BamHI E region of the Epstein Barr virus genome encodes three transformation associated nuclear proteins. *Proc. Natl. Acad. Sci. USA* 85:995-999, 1988.
215. Satterthwait, A., T. Arrhenius, R. Hagopian F. Zavala, V. Nussenzweig, and R.A. Lerner. The conformational restriction of peptidyl immunogens with covalent replacements for the hydrogen bond. *Nobel Symposium, Vaccine Vol. 6*, 99-103, April, 1988.
216. Dyson, H.J., M. Rance, R.A. Houghten, R.A. Lerner, and P.E. Wright. Folding of peptide fragments of proteins in water solution 1. Sequence requirements for the formation of a reverse turn. *Journal of Molecular Biology* 201:161-200, 1988.
217. Dyson, H.J., M. Rance, R.A. Houghten, P.E. Wright, and R.A. Lerner. Folding of peptide fragments of proteins in water solution 2. The nascent helix. *Journal of Molecular Biology* 201:201-217, 1988.
218. Tramontano, A., A. Ammann, and R.A. Lerner. Antibody catalysis approaching the activity of enzymes. *J. Am. Chem. Soc.* 110:2282-2286, 1988.
219. Field, J., J. Nikawa, D. Broek, B. MacDonald, L. Rodgers, I.A. Wilson, R.A. Lerner, and M. Wigler. Purification of a RAS responsive adenylate cyclase complex from *saccharomyces cerevisiae* using an antigen fusion-immunoaffinity method. *Mol. Cellular Biology* 8:2159-2165, 1988.
220. Benkovic, S. and R.A. Lerner. Principles of antibody catalysis. *BioEssays* 9:107-112, 1988.
221. Fieser, T., J. Tainer, M. Geysen, R.A. Houghten, and R.A. Lerner. Influence of Protein flexibility and peptide conformation on reactivity of monoclonal anti-peptide antibodies with a protein α -helix. *Proc. Natl. Acad. Sci. USA* 84:8568-8572, 1987.
222. Getzoff, E., J. Tainer, R.A. Lerner and M. Geysen. The chemistry and mechanism of antibody binding to proteins antigens. *Advances in Immunology* 43:1-98, 1988.
223. Janda, K.D., R.A. Lerner, and A. Tramontano. Antibody catalysis of bimolecular amide formation. *J. Am. Chem. Soc.* 110:4835-4837, 1988.
224. Benkovic, S.J., A.D. Napper, and R.A. Lerner. Catalysis of a stereospecific bimolecular reaction by an antibody. *Proc. Natl. Acad. Sci. USA* 85:5355-5358, 1988.
225. Wright, E.P., H.J. Dyson, and R.A. Lerner. Conformation of Peptide Fragments of Proteins in Aqueous Solution: Implications for Initiation of Protein Folding. *Biochemistry* 27:7167-7175, 1988.

THIS PAGE BLANK (USPTO)

226. Janda, K.D., D. Schloeder, S.J. Benkovic, and R.A. Lerner. Induction of an antibody which catalyzes hydrolysis of an amide bond. *Science* 241:1188-1191, 1988.
227. Wright, P.E., J.H. Dyson and R.A. Lerner. Folding of peptide fragments of proteins in water solution. "Protein Folding. Deciphering the Second Half of the Genetic Code", Eds. J.A. King and L.M. Gierasch AAAS pp. 95-102, 1988.
228. Waltho, J.P., R.A. Lerner, and P.E. Wright. Conformation of a T cell stimulating peptide in aqueous solution. *FEBS Lett.* 250:400-404, 1989.
229. Ebina, S., R.A. Lerner, and P.E. Wright. Chemical modification of bovine pancreatic trypsin inhibitor for single site coupling of immunogenic peptides for NMR conformational analysis. *J. Biol. Chem.* 264:7882-7888, 1989.
230. Tsang, P., T.M. Fieser, J.M. Ostresh, R.A. Lerner, and P. Wright. Isotope-edited NMR studies of fab'-peptide complexes. *Peptide Research* 1:1-5, 1988.
231. Iverson, B. and R.A. Lerner. Sequence specific peptide cleavage catalyzed by an antibody. *Science* 243:1184-1188, 1989.
232. Janda, K.D., S.J. Benkovic, and R.A. Lerner. Catalytic antibodies with lipase activity and R and S substrate selectivity. *Science* 244:437-440, 1989
233. Schwabacher, A.W., M.I. Weinhouse, M.T.M. Auditor, and R.A. Lerner. Metalloselective anti-porphyrin monoclonal antibodies. *J. Am. Chem. Soc.* 111:2344-2346, 1989
234. Tsang, P., T.M. Fieser, J.M. Ostresh, R.A. Houghten, R.A. Lerner, and P.E. Wright. Solution NMR studies of FAB'-peptide complexes. *Frontiers of NMR in Molecular Biology* 63-73, ed. Alan R. Liss, Inc. 1990.
235. Wright, P.E., H.J. Dyson, V.A. Feher, L.L. Tennant, J.P. Waltho, R.A. Lerner, and D.A. Case. Folding of peptide fragments of proteins in aqueous solution. "Frontiers of NMR in Molecular Biology", UCLA Symposium on Molecular and Cellular Biology, New Series, Eds. D. Live, I. Armitage, and D. Patel, Vol 109, pp. 1-13, Alan R. Liss, Inc., New York, 1990.
236. Sastry, L., M. Alting-Mees, W.D. Huse, J. M. Short, J.A. Sorge, B.N. Hay, K.D. Janda, S.J. Benkovic, and R.A. Lerner. Cloning of the immunological repertoire in *Escherichia Coli* for generation of monoclonal catalytic antibodies: Construction of a heavy chain variable region-specific cDNA library. *Proc. Natl. Acad. Sci. USA* 86:5728-5732, 1989.
237. Norrby, E., E.D. Parks, G. Utter, R.A. Houghten, and R.A. Lerner. Immunochemistry of the dominating antigenic region Ala⁵⁸² to Cys⁶⁰⁴ in the transmembranous protein of simian and human immunodeficiency virus. *J. Immunol.* 143:3602-3608, 1989.

THIS PAGE BLANK (USPTO)

238. Grove, A. and R.A. Lerner. Catalytic antibodies the effect of binding interactions in catalysis. Nordisk Insulin Symposium No. 2, of Molecular Mimicry in Health and Disease, 1988.
239. Dillner, J., L. Dillner, J. Robb, J. Willems, I. Jones, W. Lancater, R. Smith and R.A. Lerner. A synthetic peptide defines a serologic IgA response to a human papillomavirus-encoded nuclear antigen expressed in virus-carrying cervical neoplasia. *Proc. Natl. Acad. Sci, USA* 86:3838-3841, 1989.
240. Norrby, E., G. Biberfeld, P.R. Johnson, E.D. Parks, R.A. Houghten, and R.A. Lerner. The chemistry of site-directed serology for HIV infections. *AIDS Res. Human Retrov.* 5:473-479, 1989.
241. Stura, E.A., R.L. Stanfield, T.M. Fieser, R.S. Balderas, L.R. Smith, R.A. Lerner, and I.A. Wilson. Preliminary crystallographic data and primary sequence for an anti-peptide Fab' and its complex with the C-helix peptide from myohemerythrin. *J. Biological Chemistry* 264:15721-15725, 1989.
242. Wright, P.E., R.A. Lerner, and J.H. Dyson. Folding of peptide fragments of proteins in water solution: Implications for initiation of protein folding. *Advances in Protein Design* ed. H. Blocker, J. Collins, R.D. Schmid, D. Schomburg, VCH Publishing, pp. 13-19, 1989.
243. Stopner, A., G. Utter, M.A. Mufson, C. Orvell, R.A. Lerner, and E. Norrby. A subgroup-specific antigenic site in the G protein of respiratory syncytial virus forms a disulfide-bonded loop. *J. Virology* 64: October, 1990.
244. Satterthwait, A.C., T. Arrhenius, R.A. Hogopian, F. Zavala, V. Nussenzweig, and R.A. Lerner. The conformational restriction of synthetic peptides, including a malaria peptide, for use as immunogens. *Phil. Trans. R. Soc. Lond. B* 323:565-572, 1989.
245. Wright, P.W., H.J. Dyson, R.A. Lerner, L. Riechmann, and P. Tsang. Antigen-Antibody interactions: An NMR approach. *Biochemical Pharmacology* 40:83-88, 1990.
246. Iverson, S.A., L. Sastry, W.D. Huse, J.A. Sorge, S.B. Benkovic, and R.A. Lerner. A combinatorial system for cloning and expressing the catalytic antibody repertoire in *E. coli*. *CSH Symp. on Quantitative Biology, Vol. LIV.* pp. 273-281, 1989.
247. Alexander, H., Alexander, S., Heffron, F., Fieser, T.M., Hay, B.N., Getzoff, E.D., Tainer, J.A., and Lerner, R.A. Synthesis and Characterization of a recombinant myohemerythrin protein encoded by a synthetic gene. *Gene* 99:151-156, 1991.
248. Janda, K.D., M.I. Weinhouse, D.M. Schloeder, S.J. Benkovic, and R.A. Lerner. Bait and switch strategy for obtaining catalytic antibodies with acyl-transfer capabilities. *J. Am. Chem. Soc.*, 112:1274, 1989.

THIS PAGE BLANK (USPTO)

249. Huse, W.D., L. Sastry, S.A. Iverson, A.S. Kang, M. Alting-Mees, D.R. Burton, S.J. Benkovic, and R.A. Lerner. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science* 246:1275-1281, 1989.
250. Stanfield, R.L, T.M. Fieser, R.A. Lerner, and I.A. Wilson. Crystal structures of an antibody and its complex with peptide antigen at 2.8Å. *Science* 248:712-719, 1990.
251. Dyson, J.H., A.C. Satterthwait, R.A. Lerner, and P.E. Wright. Conformational Preferences of synthetic peptides derived from the immunodominant site of the circumsporozoite protein of plasmodium falciparum by ¹H NMR. *Biochemistry* 29:7828-7837, 1990.
252. Schultz, P.G., R.A. Lerner, and S.J. Benkovic. Catalytic antibodies. *Chemical and Engin. News*, 68:26-40, 1990.
253. Persson, M.A.A. and R.A. Lerner. Increased yield of antibody-producing murine spleen cell hybridomas from fusions cultured in medium supplemented with mouse serum. *J. Immunological Methods*, 127:39-42, 1990.
254. Lerner, R.A. and S.J. Benkovic. Observations in the interface between immunology and chemistry. *Chemtracts-Organic Chemistry*, 3:1-36, 1990.
255. Iverson, B.L., S.A. Iverson, V.A. Roberts, E.D. Getzoff, J.A. Tainer, S.J. Benkovic and R.A. Lerner. Metalloantibodies. *Science* 249:659-662, 1990.
256. Roberts, V.A., B.L. Iverson, S.A. Iverson, S.J. Benkovic, R.A. Lerner, E.D. Getzoff and J.A. Tainer. Antibody remodeling: A general solution to the design of a metal-coordination site in an antibody binding pocket. *Proc. Natl. Acad. Sci. USA* 87:6654-6658, 1990.
257. Benkovic, S.J., J.A. Adams, C.L. Borders, Jr., K.D. Janda, and R.A. Lerner. The enzymic nature of antibody catalysis: Evolution of multistep kinetic processing. *Science* 250:1135-1138, 1990.
258. Janda, K.D., J.A. Ashley, T.M. Jones, D.A. McLeod, D.M. Schloeder, M.I. Weinhouse, R.A. Lerner, R.A. Gibbs, P.A. Benkovic, R. Hilhorst and S.J. Benkovic. Catalytic antibodies with acyl transfer capabilities: Mechanistic and kinetic investigations. *J. Am. Chem. Soc.* 113:291-297, 1991.
259. Janda, K.D., S.J. Benkovic, D.A. McLeod, D.M. Schloeder and R.A. Lerner. Substrate Attenuation: An approach to improve antibody catalysis. *Tetrahedron* 47:2503-2506, 1991.
260. Lerner, R.A., S.J. Benkovic and P.G. Schultz. At the crossroads of chemistry and immunology: Catalytic Antibodies. *Science* 252:659-667, 1991.

THIS PAGE BLANK (USPTO)

261. Dyson, H.J., E. Norrby, M. Feinstein, D.E. Parks, R.A. Lerner and P.E. Wright. Immunogenic peptides corresponding to the dominant antigenic region Ala⁵⁹⁷ to Cys⁶¹⁹ in the transmembranous protein of simian immunodeficiency virus have a propensity to fold in aqueous solution. *Biochemistry* Vol. 31, No. 5 pp. 1458-1463, 1992.
262. Kang, A.S., C.F. Barbas, K.D. Janda, S.J. Benkovic and R.A. Lerner. Linkage of recognition and replication functions by assembling combinatorial antibody fab libraries along phage surfaces. *Proc. Natl. Acad. Sci. USA* 88:4363-4366, 1991.
263. Wirsching, P., J.A. Ashley, S.J. Benkovic, K.D. Janda and R.A. Lerner. An Unexpectedly efficient catalytic antibody operating by ping-pong and induced fit mechanisms. *Science* 252:680-685, 1991.
264. Ikeka, S., M.I. Weinhouse, K.D. Janda, R.A. Lerner and S.J. Danishefsky. Asymmetric induction via a catalytic antibody. *J. Am. Chem. Soc.* 113:7763-7764, 1991.
265. Fujii, I., R.A. Lerner, and K.D. Janda. Enantiofacial protonation by catalytic antibodies. *J. Am. Chem. Soc.* 113:8528-8529, 1991.
266. Barbas, C., A.S. Kang, R.A. Lerner and S.J. Benkovic. Assembly of combinatorial antibody libraries on phage surfaces: The gene III site. *Proc. Natl. Acad. Sci. USA*, 88:7978-7982, 1991.
267. Tsang, P., M. Rance, T.M. Fieser, J.M. Ostresh, R.A. Houghten, R.A. Lerner and P.E. Wright. Conformation and Dynamics of an Fab'-bound peptide by isotope-edited NMR spectroscopy. *Biochemistry* 31:3862-3871, 1992.
268. Benkovic, S.J., J. Adams, K.D. Janda and R.A. Lerner. A catalytic antibody uses a multistep kinetic sequence. *Ciba Foundation Symposium* 159, Wiley, Chichester, pp. 4-12, 1991.
269. Sastry, L., M. Mubarak, K.D. Janda, S.J. Benkovic and R.A. Lerner. Screening combinatorial antibody libraries for catalytic acyl transfer reaction. *Ciba Foundation Symposium* 159, Wiley, Chichester, pp. 145-155, 1991.
270. Iverson, B.L., S.A. Iverson, K.E. Cameron, G.K. Jahangiri, D.S. Pasternak and R.A. Lerner. Tritylase antibodies. *Ciba Foundation Symposium* 159, Wiley, Chichester, pp. 227-235, 1991.
271. Burton, D.R., C.F. Barbas III, M.A.A. Persson, S. Koenig, R.M. Chanock and R.A. Lerner. A large array of human monoclonal antibodies to HIV-1 from combinatorial libraries of asymptomatic seropositive individuals. *Proc. Natl. Acad. Sci. USA*, 88:10134-10137, 1991.

THIS PAGE BLANK (USPTO)

272. Reymond, J.L., K.D. Janda and R.A. Lerner. Antibody catalysis of glycosidic bond hydrolysis. *Angewandte Chemie Int, Ed. Engl.* Vol. 30 No. 12 pp. 1711-1713, 1991.
273. Lerner, R.A., C.F. Barbas, III, A.S. Kang and D.R. Burton. On the use of combinatorial antibody libraries to clone the "Fossil Record" of an individual's immune response. *Proc. Natl. Acad. Sci. USA*, 88:9705-9706, 1991.
274. Gibbs, R.A., P.A. Benkovic, K.D. Janda, R.A. Lerner and S.J. Benkovic. Substituent effects on an antibody-catalyzed hydrolysis of phenyl esters: Further evidence for an acyl-antibody intermediate. *J. Am. Chem. Soc.*, 114:3528-3534, 1992.
275. Kang, A.S., D.R. Burton and R.A. Lerner. Combinatorial immunoglobulin libraries in phage. *Methods Vol. 2, No. 2*, 111-118, (ed: Lerner and Burton) 1991.
276. Dyson, J.H., G. Merutka, J.P. Waltho, R.A. Lerner, P.E. Wright. Folding of peptides fragments comprising the complete sequence of proteins: Models for initiation of protein folding. I. Myohemerythrin. *J. Molecular Biology*, Submitted, 1991.
277. Zebedee, S.L., C.F. Barbas, III, Y.-L. Hom, R.H. Caothein, R. LaPolla, D.R. Burton, R.A. Lerner and G.B. Thornton. Human combinatorial libraries to hepatitis B surface antigen: Expression on the phage surface. *CSH Symposium*, 1991.
278. Barbas III, C.F and Richard Lerner. Combinatorial immunoglobulin libraries on the surface of phage (Phabs): Rapid selection of antigen-specific Fabs. *Methods Vol. 2, No. 2*, 119-124, (ed: Lerner and Burton) 1991.
279. Reymond, Jean-Louis, K.D. Janda and R.A. Lerner. Highly Enantioselective protonation catalyzed by an antibody. *J. Am. Chem. Soc.*, 114:2257-2258, 1992.
280. Dyson, H.J, J. Sayre, G. Merutka, H.C. Shin, R.A. Lerner and P.E. Wright. Folding of peptide fragments comprising the complete sequence of proteins: Models for initiation of protein folding II. Plastocyanin. Submitted, December 1991.
281. Alexander, H., S. Alexander, E. Getzoff, J.A. Tainer, H.M. Geysen and R.A. Lerner. Altering the antigenicity of proteins. *Proc. Natl. Acad. Sci. USA*, 89:3352-3356, 1991.
282. Zebedee, S.L., C.F. Barbas III, Y.-L. Hom, R.H. Coathien, R. LaPolla, D.R. Burton, R.A. Lerner, and G.B. Thornton. Human combinatorial antibody libraries to hepatitis B surface antigen. *Proc. Natl. Acad. Sci. USA*, 89:3175-3179, 1991.
283. Gram, Hermann, L.A. Marconi, C.F. Barbas III, T.A. Collet, R.A. Lerner and A.S. Kang. In vitro selection and affinity maturation of antibodies from a naive combinatorial immunoglobulin library. *Proc. Natl. Acad. Sci. USA*, 89:3576-3580, 1992.

THIS PAGE BLANK (USPTO)

284. Lerner, R.A. and K.D. Janda. Abzymes and Enzymes. The Chem. Soc. of Japan 52-56, 1991.
285. Barbas, III, C.F., J.D. Bain, D.M. Hoekstra and R.A. Lerner. Semi-synthetic combinatorial antibody libraries: A chemical solution to the diversity problem. Proc. Natl. Acad. Sci. USA, 89:4457-4461, 1992
286. Brenner, S., R.A. Lerner. Encoded combinatorial chemistry. Proc. Natl. Acad. Sci. USA, 89:5381-5383, 1992.
287. Fernholz, E., D. Schloeder, K. K.-C. Liu, C.W. Bradshaw, H. Huang, K. Janda, C.H. Wong and R.A. Lerner. Specificity of antibody-catalyzed transesterifications using enol esters: A comparison with lipase reactions. J. Organic Chem., 57:4756-4761, 1992.
288. Wade, W., J.S. Koh, N. Han, D.M. Hoekstra and R.A. Lerner. Engineering metal coordination sites into the antibody light chain. J. Am. Chem. Soc. 115:4449-4456, 1993.
289. Light, J. and R.A. Lerner. Phophabs: Antibody-phage alkaline phosphatase conjugates for one step Elisass without immunization. Bioorganic & Medicinal Chemistry Letters, Vol. 2, No. 9, 1073-1078, 1992.
290. Collet, T.A., P. Roben, R.O=Kennedy, d.F. Barbas, III, D.R. Burton and R.A. Lerner. A binary plasmid system for shuffling combinatorial antibody libraries. Proc. Natl. Acad. Sci. USA, 89:10026-10030, 1992.
291. Janda, K.D., C.G. Shevlin and R.A. Lerner. Antibody catalysis of a disfavored chemical transformation. Science, 259:490-493, 1993.
292. Y.-C. Jack Chen, T. Danon, L. Sastry, M. Mubarak, K.D. Janda and R.A. Lerner. Catalytic Antibodies from Combinatorial libraries. J. Am. Chem. Soc. 115:357-358, 1993.
293. Reymond, J.-L., G.K. Jahanghiri, C. Stoudt and R.A. Lerner. Antibody catalyzed hydrolysis of enol ethers. J. Am. Chem. Soc. 115:3909-3917, 1993.
294. Sarvetnick, N., D. Gurushantahaiah, N. Han, J. Prudent, P. Schultz and R.A. Lerner. Increasing the chemical potential of the germ-line antibody repertoire. Proc. Natl. Acad. Sci. USA 90:4008-4011, 1993.
295. Lerner, R.A., A.S. Kang, J.D. Bain, D.R. Burton and C.F. Barbas, III. Antibodies without immunization. Science, 258:1313-1314, 1992.
296. Wade, W.S., J.A. Ashley, G.K. Jahangiri, G. McElhaney, K.D. Janda and R.A. Lerner. A highly specific metal activated catalytic antibody. J. Am. Chem. Soc., 115:4906-4907, 1993.

THIS PAGE BLANK (USPTO)

297. Barbas, III, C.F., J.S. Rosenblum and R.A. Lerner. Direct selection of antibodies that coordinate metals from semisynthetic combinatorial libraries. *Proc. Natl. Acad. Sci. USA*, 90:6385-6389, 1993.
298. Na, J., K.N. Houk, C.G. Shevlin, K.D. Janda and R.A. Lerner. The energetic advantage of 5-EXO versus 6-ENDO epoxide openings: A preference overwhelmed by antibody catalysis. *J. Am. Chem. Soc.*, 115:8453-8454, 1993.
299. Lerner, R.A. and P.G. Schultz. Antibody catalysis of difficult chemical transformations. *Acc. Chem. Res.* 26:391-395, 1993.
300. Barbas, C.F., III, W. Amberg, A. Simoncsits, T.M. Jones and R.A. Lerner. Selection of human anti-hapten antibodies from semisynthetic libraries. *Gene* 137:57-62, 1993.
301. Gouverneur, V.E., K.N. Houk, B.de Pascual-Teresa, B. Beno, K.D. Janda and R.A. Lerner. Control of the exo and endo pathways of the Diels-Alder Reaction by antibody catalysis. *Science*, 262:204-208, 1993.
302. Reymond, J.-L., J.-L. Reber and R.A. Lerner. Multigram scale synthesis using a catalytic antibody. *Angewandte Chemie Int, Ed. Engl.*, 33. No. 4: 475-477, 1994.
303. Janda, K.D., C.-H.L. Lo, T. Li, C.F. Barbas, III, P. Wirsching and R.A. Lerner. Direct selection for a catalytic mechanism from combinatorial antibody libraries. *Proc. Natl. Acad. Sci., USA*, 91:2532-2536, 1994.
304. Koch, A., J.-L. Reymond and R.A. Lerner. Antibody-catalyzed activation of unfunctionalized olefins for highly enantioselective asymmetric epoxidation. *J. Am. Chem. Soc.*, 116:803-804, 1993.
305. Li, T., K.D. Janda, J.A. Ashley and R.A. Lerner. Antibody catalyzed cationic cyclization. *Science*, 264:1289-1293, 1994
306. Cravatte, B.F., J.A. Ashley, K.D. Janda, D.L. Boger and R.A. Lerner. Crossing extreme mechanistic barriers by antibody catalysis: Syn elimination to a Cis olefin. *J. Am. Chem. Soc.*, 116:6013-6014, 1994
307. Chen, Y., J.-L. Reymond and R.A. Lerner. An antibody-catalyzed 1,2- rearrangement of carbon-carbon bonds. *Angewandte Chemie Int, Ed. Engl.*, 33. No. 15/16, 1607-1609, 1994.
308. Lerner, R.A., G. Siuzdak, O. Prospero-Garcia, S. J. Henriksen, D.L. Boger and B. Cravatt. Cerebrodiene: A brain lipid isolated from sleep deprived cats. *Proc. Natl. Acad. Sci. USA*, 91:9505-9508, 1994.

THIS PAGE BLANK (USPTO)

309. Boger, D.L., R.A. Lerner, and B. Cravatt. Synthesis of a functionalized rigid bicyclo [2.2.1] heptane: A useful hapten for eliciting catalytic antibodies. *J. Organic Chem.*, 59:5078-5079, 1994.
310. Janda, K.D., C.G. Shevlin and R.A. Lerner. Oxepane synthesis along a disfavored pathway: The rerouting of a chemical reaction using a catalytic antibody. *J. Am. Chem. Soc.*, 117:2659-2660, 1995.
311. Li, T., K.D. Janda, S. Hilton and R.A. Lerner. An antibody catalyzed nucleophilic substitution reaction at a primary carbon that appears to proceed by an ionization mechanism. *J. Am. Chem. Soc.*, 117:2367-2368, 1995.
312. Fenniri, H., K.D. Janda and R.A. Lerner. Encoded reaction cassette for the highly sensitive detection of the making and breaking of chemical bonds. *Proc. Natl. Acad. Sci. USA*, 92:2278-2282, 1995.
313. Cravatt, B.F., O. Prospero-Garcia, G. Siuzdak, N. B. Gilula, S.J. Henrikson, D.L. Boger and R.A. Lerner. Chemical characterization of a family of brain lipids that induce sleep. *Science*, 268:1506-1509, 1995.
314. Light, J. and R.A. Lerner. Random mutagenesis of staphylococcal nuclease and phage display selection. *Bioorg. & Med. Chem.* Vol. 3, No. 7, 955-967, 1995.
315. Rosenblum, J.S., L.-C. Lo, T. Li, K.D. Janda and R.A. Lerner. Antibody-catalyzed phosphate triesters hydrolysis. *Angewandte Chemie Int, Ed. Engl.*, 34, No. 20, 2275-2277, 1995.
316. Koch, T., J.-L. Reymond and R.A. Lerner. Antibody catalysis of multistep reactions: an Aldol addition followed by a disfavored elimination. *J. Am. Chem. Soc.*, 117:9383-9387, 1995.
317. Schultz, P.G., and R.A. Lerner. From molecular diversity to catalysis: Lessons from the Immune system. *Science*, 269: 1835-1842, 1995.
318. Wagner, J., R.A. Lerner, C.F. Barbas. Efficient aldolase catalytic antibodies that use the enamine mechanism of natural enzymes. *Science*, 270:1797-1800, 1995.
319. Wirsching, P., J.A. Ashley, C.-H. L. Lo, K.D. Janda and R.A. Lerner. Reactive Immunization. *Science*, 270:1775-1782, 1995.
320. Cravatt, B.F., R.A. Lerner, D.L. Boger. Structure determination of an endogenous sleep-inducing lipid, cis-9-octadecenoamide (oleylamide): A synthetic approach to the chemical analysis of trace quantities of a natural product. *J. Am. Chem. Soc.*, 118:580-590, 1996

THIS PAGE BLANK (USPTO)

321. Li, T., K.D. Janda, R.A. Lerner. Cationic Cyclopropanation by antibody catalysis. *Nature*, 379:326-327, 1996.
322. Lerner, R.A. and C.F. Barbas, III. Using the process of reactive immunization to induce catalytic antibodies with complex mechanisms: Aldolases. *Acta Chemica Scandinavica*, 50:672-678, 1996.
323. Bahr, N., R. Guller, J.L. Reymond and R.A. Lerner. A nitroaryl synthase catalytic antibody. *J. Am. Chem. Soc.*, 118:3550-3555, 1996.
324. Patterson, J.E., I.R. Ollmann, B.F. Cravatt, D.L. Boger, C.H. Wong and R.A. Lerner. Inhibition of oleamide hydrolase catalyzed hydrolysis of the endogenous sleep-inducing lipid cis-9-octadecenamide. *J. Am. Chem. Soc.*, 118:5938-5945, 1996.
325. Hasseroth, J., K.D. Janda and R.A. Lerner. Antibody catalyzed terpenoid cyclization. *J. Am. Chem. Soc.*, 118:11654-11655, 1996.
326. Lerner, R.A. and K.D. Janda. Catalytic antibodies: Evolution of protein function in real time. *Interface Between Chemistry and Biochemistry*, ed. By P. Jolles and H. Jornvall, Birkhauser Verlag Basel/Switzerland, 121-138, 1995.
327. Rosenblum, J.S., N.B. Gilula and R.A. Lerner. On signal sequence polymorphisms and diseases of distribution. *Proc. Natl. Acad. Sci, USA*, 93:4471-4473, 1996.
328. Bjornestedt, R., G. Zhong, R.A. Lerner and C.F. Barbas, III. Copying nature's mechanism for the decarboxylation of β -ketoacids into catalytic antibodies by reactive immunization. *J. Am. Chem. Soc.*, 118:11720-11724, 1996.
329. Cravatt, B.F., D.K. Giang, S.P. Mayfield, D.L. Boger, R.A. Lerner and N.B. Gilula. Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amide. *Nature*, 384:83-87, 1996.
330. Wagner, J., R.A. Lerner and C.F. Barbas, III. Synthesis of five enantiomerically pure haptens designed for in vitro evolution of antibodies with peptidase activity. *Bioorganic & Medicinal Chemistry*, 4:901-916, 1996.
331. Keinan, E. and R.A. Lerner. The first decade of antibody catalysis: Perspective and prospects. *Israel Journal of Chemistry*, 36:113-119, 1996.
332. Lo, C.-H.L., C. Gao, S. Mao, K. Matsui, R. A. Lerner and K.D. Janda. Chain shuffling: Investigations into the specificity and selectivity of antibody catalysis. *Israel Journal of Chemistry*, 36:195-198, 1996.

THIS PAGE BLANK (USPTO)

333. Hasserodt, J., K.D. Janda and R.A. Lerner. Formation of bridge-methylated decalins by antibody-catalyzed tandem cationic cyclization. *J. Am. Chem. Soc.*, 119:5993-5998, 1997.
334. Janda, K.D., L.-C. Lo, C.-H. Lo, M.M. Sim, C.-H. Wong and R.A. Lerner. Chemical selection for catalysis in combinatorial antibody libraries. *Science*, 275:945-948, 1997.
335. Li, T., R.A. Lerner and K.D. Janda. Antibody-catalyzed cationic reactions: Rerouting of chemical Transformation via antibody catalysis. *Acc. Chem. Res.* 30:115-117, 1997.
336. Zhong, G., T. Hoffmann, R.A. Lerner, S. Danishefsky and C.F. Barbas, III. Antibody catalyzed enantioselective Robinson annulation. *J. Am. Chem. Soc.*, 119:8131-8132, 1997.
337. Lerner, R.A. A hypothesis about the endogenous analogues of general anesthesia. *Proc. Natl. Acad. Sci., USA*, 94:13375-13377, 1997.
338. Boger, D.L., J.E. Patterson, X. Guan, B.F. Cravatt, R.A. Lerner and N.B. Gilula. Chemical requirements for inhibition of gap junction communication by the biologically active lipid oleamide. *Proc. Natl. Acad. Sci., USA*, 95:4810-4815, 1998.
339. Schroer, J., M. Sanner, J.-L. Reymond and R.A. Lerner. Design and synthesis of transition state analogs for induction of hydride transfer catalytic antibodies. *J. Org. Chem.* 62:3220-3229, 1997.
340. Lin, C.-H., T. Z. Hoffman, P. Wirsching, C.F. Barbas, III, K.D. Janda and R.A. Lerner. On roads not taken in the evolution of protein catalysts: Antibody steroid isomerases that use an enamine mechanism. *Proc. Natl. Acad. Sci., USA*, 94:11773-11776, 1997.
341. Gao, C., C.-H. Lin, C.-H. Lo, S. Mao, P. Wirsching, R.A. Lerner and K.D. Janda. Making chemistry selectable by linking it to infectivity. *Proc. Natl. Acad. Sci., USA*, 94:11777-11782, 1997.
342. Barbas, III, C.F., A. Heine, G. Zhong, T. Hoffmann, S. Gramatikova, R. Bjornestedt, B. List, J. Anderson, E.A., Stura, I.A. Wilson and R. A. Lerner. Immune versus natural selection: Antibody aldolases with enzymic rates but broader scope. *Science*, 278:2085-2098, 1997.
343. List, B., D. Shabat, C.F. Barbas, III and R.A. Lerner. Enantioselective total synthesis of some brevicomins using aldolase antibody 38C2. *Chem. Eur. J.*, 4:881-885, 1998.
344. Guan, X., B.F. Cravatt, G.R. Ehrling, J.E. Hall, D.L. Boger, R.A. Lerner and N.B. Gilula. The sleep-inducing lipid oleamide deconvolutes gap junction communication and calcium wave transmission in Glial Cells. *The Journal of Cell Biology*, 139:1-8, 1997.

THIS PAGE BLANK (USPTO

345. Hoffmann, T., G. Zhong, B. List, D. Shabat, J. Anderson, S. Gramatikova, R.A. Lerner and C.F. Barbas, III. Aldolase antibodies of remarkable scope. *J. Am. Chem. Soc.*, 120:2768-2779, 1998.
346. Finn, M.G., Lerner, R.A. and C.F. Barbas, III. Cofactor-induced refinement of catalytic antibody activity: A metal-specific allosteric effect. *J. Am. Chem. Soc.*, 120:2963-2964, 1998.
347. Zhong, G., D. Shabat, B. List, J. Anderson, S.C. Sinha, R.A. Lerner and C.F. Barbas, III. Catalytic enantioselective retro-aldol reaction: Kinetic resolution of β -hydroxyketones with aldolase antibodies. *Angewandte Chemie Int, Ed. Engl.*, 37, No. 18, 2481-2484, 1998.
348. Taylor, M.J., T.Z. Hoffman, J.T. Yli-Kauhaluoma, R.A. Lerner and K.D. Janda. A light-activated antibody catalyst. *J. Am. Chem. Soc.*, 120:12783-12790, 1998.
349. Sinha, S., C., C.F. Barbas and R.A. Lerner. The antibody catalysis route to the total synthesis of epothilones. *Proc. Natl. Acad. Sci. USA*, 95:14603-14608, 1998.
350. List, Benjamin, C. F. Barbas III and R.A. Lerner. Aldol sensors for the rapid generation of turnable fluorescence by antibody catalysis. *Proc. Natl. Acad. Sci. USA*, 95:15351-15355, 1998.
351. Lerner, Richard, C.F. Barbas III, and K.D. Janda. Making Enzymes. The Harvey Lecture Series, 92:1-40, 1998
352. Grubber, K., B. Zhou, K.N. Houk, R.A. Lerner, C.G. Shevlin and I.A. Wilson. Structural basis for antibody catalysis of a disfavored ring closure reaction. *Biochemistry*, 38:7062-7074, 1999.
353. Shabat, D., B. List, R.A. Lerner and C.F. Barbas III. A short enantioselective synthesis of 1-Deoxy-L-xylulose by antibody catalysis. *Tetrahedron Letters*, 40:1437-1440, 1999.
354. Gao, Changshou, Shenlan Mao, C.-H. L. Lo, P. Wirsching, R. A. Lerner and K.D. Janda. Making artificial antibodies: A new format for phage display of combinatorial heterodimeric arrays. *Proc. Natl. Acad. Sci. USA*, 96:6025-6030, 1999.
355. Shabat, D., C. Rader, B. List, R.A. Lerner and C.F. Barbas, III. Multiple event activation of a generic produrg trigger by antibody catalysis. *Proc. Natl. Acad. Sci. USA*, 96:6925-6930, 1999.
356. Zhong, G., R.A. Lerner and C.F. Barbas, III. Broadening the aldolase catalytic antibody repertoire by combining reactive immunization and transition state theory: New enantio- and diastereoselectivities., *Angew. Chem.*, 38:3738-3741, 1999.

THIS PAGE BLANK (USPTO)

357. Paschall, C.M., J. Hasserodt, T. Jones, R.A. Lerner, K.D. Janda and D.W. Christianson. Convergence of catalytic antibody and terpene cyclase mechanisms: Polyene cyclization directed by carbocation- π interactions. *Angew. Chem. Int. Ed.*, 38:1743-1747, 1999.
358. List, B., R.A. Lerner and C.F. Barbas, III. Enantioselective aldol cyclodehydrations catalyzed by antibody 38C2. *Organic Letters*, 1:59-61, 1999.
359. List, B., D. Shabat, G. Zhong, J.M. Turner, A. Li, T. Bui, J. Anderson, R.A. Lerner and C.F. Barbas, III. A catalytic enantioselective route to hydroxy-substituted quaternary carbon centers: Resolution of tertiary aldols with a catalytic antibody. *J. Am. Chem. Society*, 121:7283-7291, 1999.
360. Lerner, R.A., R.J. Glasscock and F.J. Dixon. The role of anti-glomerular basement membrane antibody in the pathogenesis of human glomerulonephritis. *J. Am. Soc. Nephrology*, 10:1389-1404, 1999.
361. Tanaka, F., R.A. Lerner and C.F. Barbas, III. Thiazolium-dependent catalytic antibodies produced using a covalent modification strategy. *Chemistry Communication*, 1383-1384, 1999.
362. Tanaka, F., H. Almer, R.A. Lerner and C.F. Barbas, III. Catalytic single-chain antibody possessing β -lactamase activity selected from a phage displayed combinatorial library using a mechanism-based inhibitor. *Tetrahedron Letters*, 40:8063-8066, 1999.
363. Sinha, S., J. Sun, G. Miller, C.F. Barbas, III and R.A. Lerner. Sets of aldolase antibodies with antipodal reactivities. Formal synthesis of epothilone E by large-scale antibody-catalyzed resolution of thiazole aldol. *Organic Letters*, 1:1623-1626, 1999.
364. Ly, D., D.J. Lockhart, R.A. Lerner and P.G. Schultz. Mitotic misregulation and human aging. *Science*, 287:2486-2492, 2000.
365. Karlstrom, A., G. Zhong, C. Rader, N. Larsen, A. Heine, R. Fuller, B. List, F. Tanaka, I.A. Wilson, C.F. Barbas, III and R.A. Lerner. Using antibody catalysis to study the outcome of multiple evolutionary trials of a chemical task. *Proc. Natl. Acad. Sci., USA*, 97:3878-3883, 2000.
366. Tanaka, F., R.A. Lerner, C.F. Barbas, III. Reconstructing Aldolase antibodies to alter their substrate specificity and turnover. *J. Am. Chem. Soc.* 122:4835-4836, 2000.
367. Simenov, A, M. Matsushita, E.A. Juban, E.H.Z. Thompson, T.Z. Hoffman, M.J. Taylor, P. Wirsching, W. Rettig, J.K. McCusker, D.P. Millar, P.G. Schultz, R.A. Lerner and K.D. Janda. Blue-Fluorescent Antibodies. *Science*, 290:307-313, 2000

THIS PAGE BLANK (USPTO)

368. Wentworth, A., L.H. Jones, P. Wentworth, Jr., K.D. Janda, R.A. Lerner. Antibodies have the intrinsic capacity to destroy antigens. *Proc. Natl. Acad. Sci., USA*, 97:10930-10935, 2000.
369. Turner, J.M., T. Bui, R.A. Lerner, C.F. Barbas, III, and B. List. An efficient benchtop system for multigram-scale kinetic resolutions using aldolase antibodies. *Chem. Eur. J.* 6:2772-2774, 2000.
370. Hasserodt, J., K.D. Janda and R.A. Lerner. Antibodies mimic natural oxidosqualene-cyclase action in steroid ring a formation. *J. Am. Chem. Soc.* 122:40-45, 2000.

THIS PAGE BLANK (USPTO)

Klaus Düring: Wound-Inducible Expression and Secretion of T4 Lysozyme and Monoclonal Antibodies in *Nicotiana tabacum*

Dissertation

Job No.: 1596-81522

Ref.: 310098.01

Translated from German by the Ralph McElroy Translation Company
910 West Avenue, Austin, Texas 78701 USA

Wound-Inducible Expression and Secretion of T4 Lysozyme and Monoclonal Antibodies
in *Nicotiana tabacum*

Inaugural Dissertation

to

receive the degree of

Doctor in the Mathematics and Sciences College
of Cologne University

Submitted by

Klaus Düring

from Cologne

Cologne 1988

Referees:

Prof. Jozef St. Schell

Prof. José A. Campos-Oriega

Prof. Fritz Kreuzaler (RWTH Aachen)

Date of oral examination: July 9, 1988

- I. Introduction
- II. Material and methods
 - 2.1 Bacterial strains
 - 2.2 Plant material
 - 2.3 Chemicals
 - 2.4 Cooperation and obtained materials
 - 2.5 Microbiological methods
 - 2.5.1 Culturing of *E. coli*
 - 2.5.2 Culturing of agrobacteria
 - 2.5.3 Conjugation between *E. coli* and *Agrobacterium tumefaciens*
 - 2.6 Plant tissue culture and transformation methods
 - 2.6.1 Culturing conditions
 - 2.6.2 Leaf disk test for regeneration of transformed plants
 - 2.6.3 Transient expression in tobacco leaf protoplasts
 - 2.6.4 Stable transformation of tobacco leaf protoplasts
 - 2.6.5 Induction of pT_R and pNOS in plant tissue
 - 2.7 Methods of recombinant DNA technology
 - 2.7.1 Isolation of plasmid DNA from *E. coli*
 - 2.7.2 Isolation of total DNA from agrobacteria
 - 2.7.3 Cloning methods
 - 2.7.4 Southern blot
 - 2.7.5 Hybridization of filter-bound plasmid and agrobacteria DNA with biotin-labeled samples
 - 2.8 Preparation of affinity-purified polyclonal antibodies against T4 lysozyme
 - 2.8.1 Immunization of rabbits with T4 lysozyme
 - 2.8.2 Enrichment of IgG fraction from blood serum
 - 2.8.3 Coupling of T4 lysozyme to CNBr-activated Sepharose 4B
 - 2.8.4 Purification of IgG fraction on protein A-Sepharose CL-4B
 - 2.8.5 Affinity purification of anti-T4 lysozyme antibodies on Sepharose 4B-coupled T4 lysozyme
 - 2.9 Transformation analysis of regenerated plant material
 - 2.9.1 Nopaline test
 - 2.9.2 Neomycin phosphotransferase II test (NPT test)
 - 2.10 DNA analysis of transformed plants by Southern blotting
 - 2.10.1 Isolation of genomic plant DNA
 - 2.10.2 Hybridization of filter-bound genomic DNA with radioactively labeled probes

- 2.11 RNA analysis of transformed plants by Northern blotting
 - 2.11.1 Isolation of total RNA from plants
 - 2.11.2 RNA formaldehyde agarose gels
 - 2.11.3 Hybridization of filter-bound RNA with radioactively labeled probes
 - 2.11.4 Isolation of total RNA from protoplasts
- 2.12 Protein analysis of transformed plants
 - 2.12.1 Tissue printing
 - 2.12.2 Isolation of total protein from plants
 - 2.12.3 Isolation of total protein from protoplasts
 - 2.12.4 Isolation of protein from protoplast culture medium
 - 2.12.5 $(\text{NH}_4)_2\text{SO}_4$ precipitation of T4 lysozyme plant extract (PBS)
 - 2.12.6 Coupling of polyclonal anti-T4 lysozyme antibodies to CNBr-activated Sepharose 4B
 - 2.12.7 Affinity purification of exogenous proteins from plant extract
 - 2.12.8 Western blotting
 - 2.12.9 Employed antibodies for immunodetection
 - 2.12.10 Silver staining of polyacrylamide gels
 - 2.12.11 T4 lysozyme activity test with *Micrococcus lysodeicticus*
- 2.13 Immunogold labeling
 - 2.13.1 High pressure freezing preparation
 - 2.13.2 Freezing substitution
 - 2.13.3 Ultramicrotomy
 - 2.13.4 Immunocytochemistry

III. Results

- 3.1 Construction of a T_R promoter-NOS promoter vector with chimeric genes for α -amylase signal peptide and light chain or heavy chain of the antibody
- 3.2 Construction of the T_R promoter vector with the chimeric gene for α -amylase signal peptide and T4 lysozyme
- 3.3 Nonradioactive labeling of DNA fragments for hybridizations by means of the biotin-streptavidin systems, development of a simple method
- 3.4 Integration of antibody construction in *Agrobacterium tumefaciens* and checking of transformants by Southern hybridization
- 3.5 Integration of lysozyme construction in *Agrobacterium tumefaciens* and checking of transformants by Southern hybridization
- 3.6 Expression analysis of p T_R and pNOS
- 3.7 Expression of lysozyme construct pSR 2-4 in *Acetabularia mediterranea*

- 3.8 Transient expression in tobacco protoplasts
- 3.9 Transformation of *Nicotiana tabacum* W38 with antibody construct and transformation analysis
- 3.10 Transformation of *Nicotiana tabacum* with lysozyme construct and transformation analysis
- 3.11 Integration analysis of antibody plants by Southern hybridization
- 3.12 Integration analysis of lysozyme plants by Southern hybridization
- 3.13 Transcription analysis by Northern blotting
- 3.14 Expression analysis of antibody plants by Western blotting
- 3.15 Expression analysis of lysozyme plants by Western blotting
- 3.16 Localization of formed antibodies by immunogold labeling
- 3.17 Localization of formed T4 lysozyme by immunogold labeling

IV. Discussion

V. Summary

VI. Literature

Abbreviations

A ₂₆₀	Absorption at a wavelength of 260 nm
Amp	Ampicillin
BAP	Benzylaminopurine
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
bp	Base pairs
BSA	Bovine serum albumin
CaMV	Cauliflower mosaic virus
cap	α -Aminocaproic acid
CAT	Chloramphenicol acetyltransferase
Cb	Carbenicillin
Ci	Curie
2,4-D	2,4-Dichlorophenoxyacetic acid
DMF	Dimethylformamide
dNTP	Nucleotide
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetate, sodium salt

ER	Endoplasmic reticulum
EtBr	Ethidium bromide
EtOH	Ethanol
HEPES	N-(2-hydroxyethyl)piperazine-N'-2-ethanesulfonic acid
Ig	Immunoglobulin
IPTG	Isopropyl- β -D-thiogalactoside
kb	Kilobase pairs
Km	Kanamycin
M	Molar
MM	Molecular weight
MES	Morpholinoethanesulfonic acid
MOPS	(3-(N-Morpholino))propanesulfonic acid
NAA	Naphthyl-1-acetic acid
NBT	Nitro blue tetrazolium
NiP	(5-Iodo-4-hydroxy-3-nitrophenyl)acetyl
NP	(4-Hydroxy-3-nitrophenyl)acetyl
NPT	Neomycin phosphotransferase II
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonyl fluoride
PPO	2,5-Diphenyloxazole
PVP	Polyvinylpyrrolidone
PVPP	Polyvinylpolypyrrolidone
Rif	Rifampicin
RT	Room temperature
SA-CIAP	Streptavidin-alkaline phosphatase conjugate
Sm	Streptomycin
Sp	Spectinomycin
Tet	Tetracycline
Tris	Tris(hydroxymethyl)aminomethane
X-Gal	5-Bromo-4-chloro-3-indolyl- β -galactopyranoside

I thank Professor Jeff Schell for offering the possibility to work at the Max Planck Institute for Cultivation Research and for his sponsorship and support of my work.

I thank Professor Kreuzaler for accepting me at his laboratory at MPI for Cultivation Research and for introduction into the working techniques of molecular biology and genetic engineering and for valuable motivation for my work.

Special thanks go to Dr. Sigrun Hippe, who conducted the electron microscopy in close cooperation with me and therefore made important contributions to the success of this work. I also thank her for countless valuable discussions and introduction to the methods of immunogold labeling.

I thank Dr. Robert Masterson for his excellent cooperation and many long discussions concerning analysis of the double promoter.

I thank Susanne Lipphardt for introduction into the working techniques of transient expression and Dr. Rudiger Hain for performing a naked DNA transfer in protoplasm.

All coworkers at the laboratory and institute are thanked for many small kindnesses, especially Inge Czaja and Ulrike Haus, who provided valuable assistance during their laboratory work.

I. Introduction

The development of genetic engineering methods has resulted in recent years in attempts to introduce new resistances to plants. Cultivation of resistant varieties previously was only possible in the conventional manner by crossing. In so doing, the fact that undesired properties could also be introduced always had to be tolerated. Another drawback is the work often lasting several decades with this cultivation requirement. In this period the requirement profiles can change so strongly that the entire project becomes untenable.

It has become possible in recent years through developed gene transfer methods to deliberately incorporate individual genes in plants for characteristic properties. In so doing, one or more properties of a variety can be predictably altered.

After initial attempts with marker genes, such genes have recently been incorporated increasingly in plants by genetic engineering transformation with the intention of usefully altering these plants. So-called resistance genes are chiefly at issue here, from which it is expected that they will impart resistance to herbicides, pests or diseases. Despite some success, there are still many years before such altered seed can become available to agriculture. The present study is to make a fundamental contribution to these problems.

Some important current research activities will be briefly characterized below (Eckes et al., 1987). One field being worked on as industrial research is herbicide resistance. Through manipulated plants with increased resistance to certain herbicides, their use is supposed to be made more economical, and by natural resistance this is only possible to a restricted extent and only for selective herbicides. The location of the effect in most modern herbicides is known and in many cases a gene has been isolated whose genetic product ensures detoxification of the herbicide. This can be achieved by altering the location of action by mutation, an increase in concentration at the site of action by introduction of metabolizing enzymes. The first published study concerned the herbicide glyphosate, which inhibits synthesis of aromatic amino acids in plants by inhibiting the enzyme 5-enolpyruvylshikimic acid-3-phosphate synthase (EPSP synthase) (Steinrücken and Amrhein, 1980). By cloning the EPSP synthase gene behind the constitutionally expressed 35S promoter and introduction of this construct to petunia cells, a 40-fold increased enzyme activity could be detected in calli, which is considered the cause for the now existing glyphosate tolerance (Shah et al., 1986). With a mutated EPSP synthase enzyme from *Salmonella typhimurium* (replacement of one of the total of 421 amino acids (Stalker et al., 1985)), glyphosate-tolerant tobacco plants were obtained (Comai et al., 1985). DeBlock et al. (1987) recently reported on artificial resistance to the herbicide phosphinotricin (PTT), an inhibitor of glutamine synthetase (Donn et al., 1984). In this case a resistance gene (phosphinotricin acetyltransferase, PAT) was isolated from *Streptomyces hygroscopicus*

(Thompson et al., 1987) and transferred to tobacco, tomatoes and potatoes. Transgenic plants that express PAT are resistant by metabolism to high doses of PTT (DeBlock et al., 1987).

The second important area is viral resistance. Significant successes have recently been achieved here. Natural reactions to a viral infection are the formation of pathogenesis-related proteins (PR proteins) or, after infection with a weakly pathogenic virus, the formation of cross resistance. In various plants after infection with viruses, viroids, bacteria or fungi, the occurrence of specific PR proteins that can be correlated with the formation of resistance to further infections was observed (Parent and Asselin, 1984; van Loon, 1985; Carr et al., 1987; van Loon et al., 1987). Synthesis of PR proteins is induced by the infection; the concentration increases by a factor of 100 (Hooft van Huijsduijnen et al., 1985, 1986). Extracellular localization was detected by vacuum infiltration (Parent and Asselin, 1984) and by immunofluorescence (Carr et al., 1987); the concentrations are highest in the vicinity of mesophyll cells and viral lesions. Some PR proteins were recently assigned chitinase or 1,3- β -glucanase activity (Kauffmann et al., 1987; Legrand et al., 1987; Shinshi et al., 1987; Kombrink et al., 1988). Hybrids that constitutively express a PR protein that demonstrates significantly increased resistance to viruses were obtained by crossing pathogen-induced tobacco varieties (Ahl and Gianinazzi, 1982). The gene sequences of some of these proteins were recently cloned (Hooft van Huijsduijnen, 1986; Cornelissen et al., 1987; Shinshi et al., 1987) and a practical evaluation of these resistance mechanisms is therefore available, analysis and use of infection-inducible promoters and efficiently operating signal peptides being of particular interest.

Starting from the observation of cross resistance, two groups attempted to express the gene for the coat protein of the tobacco mosaic virus (TMV) in tobacco (Bevan et al., 1985; Abel et al., 1986). A higher percentage of transformed plants that then formed the coat protein actually exhibited no symptoms after artificial infection with TMV (Abel et al., 1986), and expression of a single gene is obviously sufficient for formation of cross resistance. An attempt made in the meantime to prevent spreading of viruses by expression of antisense RNA failed again, since no success was found. On the other hand, the integration and expression of cDNA from viral satellite RNA which acts antagonistically on the actual virus were successful in tobacco plants (Courtice, 1987). Harrison et al. (1987) introduced satellite RNA of the cucumber mosaic virus (CMV), which is only replicated in CMV-infected cells; production of satellite RNA in transgenic plants (Baulcombe et al., 1986) is dependent on infection and largely suppresses the formation of disease symptoms, as well as multiplication of CMV. For tomato aspermy virus (TAV), a close relative of CMV, only suppression of the symptoms can be found. Similar results were obtained by Gerlach et al. (1987) for the tobacco ringspot virus (TobRV).

The first attempts have since also been made for artificial resistances against pests, for example, insects. Natural defense mechanisms of plants are probably triggered by fragments of

the cell wall (oligosaccharides) (Davis et al., 1986) and these can have an inhibiting effect on auxin synthesis (Albersheim, personal communication). A known reaction is the formation of protease inhibitors as a reaction to damage to the plants, these being proteins that act as potent inhibitors of serine proteases, which are involved in the digestive process of insects (Broadway et al., 1986). Genes for such inhibitors were recently isolated (Lee et al., 1986; Sanchez-Serrano et al., 1986). Different low-molecular substances, for example, phytoalexins and tannins, are also formed in addition to protease, which can have a antifeedant effect or act as signal substances (Darvill and Albersheim, 1984). Genetic engineering application of these mechanisms appears to be difficult in contrast to protease inhibitors because of their complexity. An artificial resistance to insects could be produced by transformation of plants with the endotoxin from two different *Bacillus thuringiensis* strains (Fischhoff et al., 1987; Vaeck et al., 1987). Vaeck et al. (1987) reduced the endotoxin to an effective protein fragment, and resistance to insect larvae was also produced with low-expressing regenerants. Fischhoff et al. (1987) also demonstrated resistances, although the endotoxin protein itself could not be directly detected in plants. A similar effect was achieved by Hilder et al. (1987) by expression of a trypsin inhibitor from cow bean (*Vigna unguiculata*), a roughly 80 amino acid-long polypeptide in tobacco.

The least progress has thus far been achieved in the field of genetic engineering treatment of plant diseases, primarily fungal infections. Natural "induced" resistances, similar to viral infection (Dean and Kuc, 1985) can lead to the formation of hydroxyproline-rich proteins (proteins localized in the cell wall with unknown functions) and other infection-specific proteins (Riggleman et al., 1985). cDNA and genomic clones for hydroxyproline-rich proteins have since been isolated (Chen and Varner, 1985a/b). Elicitors and inducible genes have since been characterized. Plants synthesize the enzyme chitinase whose activity increases sharply during fungal attack (Schlumbaum et al., 1986; Legrand et al., 1987; Shinshi et al., 1987). Two chitinase genes (from *Serratia marcescens*) have been identified and cloned, one of which clones for a secreted protein with a 23 amino acid-long signal peptide (Jones et al., 1986). This was expressed in tobacco plants under the control of a strong promoter (P. Dunsmuir, personal communication). It is assumed that chitinase hydrolyzes the fungal cell walls and in this manner can produce resistance.

The present study attempts to take up new aspects in this field and to solve the fundamental problems for additional resistance approaches. In the first place, expression and assembling of the light and heavy chains of an antibody and expression of bacteriophage T4 lysozyme in plants are to be investigated, and in the second place, the question whether it is possible to secrete these proteins under the guidance of a plant signal peptide is to be answered. Application of monoclonal antibodies is conceivable, for example, to generate artificial resistance against viruses (cf. Stieger, 1987).

A more specific background involves the use of lysozyme: in similar fashion to natural lysis, an attempt can be made to achieve artificial lysis of infecting bacteria. This aspect appears to be of particular interest, since thus far there are no plant protection agents against bacteria. In order for all these approaches to lead to greater activity with proteins not endogenous to plants, the combination with secretion was considered in order to make the attacker harmless already before reaching the cytoplasm. The possible presence of a signal peptide commonly found in plants in foreign proteins, with the correct function, cannot be expected with certainty. By fusion with a signal peptide endogenous to the plant, secretion should be made possible. Not only secretion, but also controllable expression of gene products are part of the construction of an ideal resistance gene. An intense search for appropriate promoters was made. The T_R double promoter employed here (Velten et al., 1984) proved to be an appropriate candidate during this study. Wound-induced promoters or those that can be simulated by other infectious phenomena would protect the plants from any stresses by exogenous proteins in the healthy state, but in the event of infection could ensure efficient resistance reaction.

The mechanisms of the process of secreting proteins by the bacterial plasma membrane (Benson et al., 1985; Oliver, 1985; Wu and Tai, 1986) and the eukaryotic endoplasmic reticulum (Kelly, 1985) have been very intensely investigated in recent years. Exocytosis is fundamentally similar for both systems but there are deviations in essential details. The eukaryotic secretion mechanism will be briefly sketched below. The first hypothesis (signal hypothesis) was advanced in 1975 by Blobel and Dobberstein (1975a/b). According to it, translocation of polypeptides is coupled to protein biosynthesis. Ribosomes bound to the membrane of the rough endoplasmic reticulum of eukaryotic cells produce secreted proteins, whereas free ribosomes synthesize cytoplasmic proteins (Redman and Sabatini, 1966). Nearly all secreted proteins have an N-terminal signal peptide of 20 to 40 amino acids which exhibits specific features (v. Heijne, 1985a) and is cleaved by signal peptidase, a membrane protein (Oliver, 1985). If the signal peptide leaves the ribosome during protein biosynthesis, which corresponds to the achievement of a length of about 70 amino acids of the mature polypeptide (v. Heijne, 1985b), protein synthesis is stopped by an interaction of the signal peptide with the signal recognition particle (SRP) (Walter et al., 1984; Wiedmann et al., 1987a). SRP consists of six polypeptides and a 300-bp RNA (7S-RNA) (Walter and Blobel, 1981, 1982) and interacts through the 54 kDa polypeptide with the signal peptide (Kurzchalia et al., 1985; Wiedmann et al., 1987a). The translation stop is eliminated by a direct interaction of SRP with its SRP receptor, also called docking protein (Meyer et al., 1982). This 73-kDa protein (Lauffer et al., 1985) is an integral component of the membrane of the rough endoplasmic reticulum, but consists of a membrane-anchoring and interacting cytosol subunit (Hortsch et al., 1985; Lauffer et al., 1985); the developing interaction is transient in nature and causes dissociation of the entire complex, but

does not participate in direct interaction between the ribosome and the membrane (Gilmore and Blobel, 1983), but only leads to a spatial approach (Walter et al., 1984). Other ER membrane proteins that participate in transport of secreted proteins through the ER membrane (Hortsch et al., 1986; Robinson et al., 1987) and for which signal peptide binding by the domains embedded in the membrane can be detected in two cases by photoaffinity labeling (Robinson et al., 1987) were recently discovered. Specific functions, namely binding of the signal peptide freed from the SRP-SRP receptor-ribosome complex, could be coordinated to a second receptor, the signal sequence receptor (SSR), SSR being an integral glycosylated protein of the membrane of the rough ER (Wiedmann et al., 1987b). According to the current view, this regulated system of a signal peptide and its receptors serves to maintain a tertiary structure of the corresponding protein appropriate for transport through the membrane (Zimmermann and Meyer, 1986). The first interaction of signal peptide and membrane occurs in an environment that is still accessible to aqueous reagents (Gilmore and Blobel, 1985).

The most recent hypothesis claims that, in mammalian cells, SRP binds the signal peptide and interacts with the SRP receptor, which is already an ER membrane protein. The signal peptide is transferred to SSR from this complex without protein biosynthesis having to be continued. The task of SRP therefore appears to be prevention of folding of the nascent protein (Walter 1987; Wiedmann et al., 1987b).

Post-translational secretion without involvement of the SRP-SRP receptor complex has also been observed (Walter, 1987; Wiedmann et al., 1987b), in this case the conformation of the complete protein still permitting interaction with SSR. This result correlates with the earlier observation of constitutional and post-translational secretion (Kelly, 1985). In the case of prepromellitin from the honeybee, Muller and Zimmermann (1987) were able to detect cotranslational insertion *in vitro* in dog pancreas microsomes without participation of SRP and SRP receptor, in which case the cluster of negatively charged amino acids on the N-terminus had to be compensated by a cluster of positively charged residues on the C-terminus, which also indicates the structural effect.

Signal peptides do not show extensive homologies in primary structure, but common structural features. The length is variable, ranging from 16 to more than 50 amino acids, but averages about 20 (v. Heijne, 1985a). The main feature is a hydrophobic core sequence of 8 to 18 uncharged amino acid residues (h region), which could be the objective for SRP. The N-terminus is very variable in length, but not in positive excess charge, which is always about +1.7. The C-terminal region of the signal peptide does not vary very strongly in length, but is polar (v. Heijne, 1985a), mostly negative in overall charge (v. Heijne, 1986a). The spatial structure (Robinson and Austen, 1987) and the charge differences between the individual domains (v. Heijne, 1985b) must be considered the main features of a signal peptide. In an *in vitro*

experiment, Briggs et al. (1986) were able to demonstrate that a synthetic signal peptide is randomly coiled in aqueous solution, and on contact with a lipid bilayer is preferably in a β -folded sheet and in inserted form in the α -helix structure. Calculations of the secondary structures also lead to a dominance of the α -helix in the h region (Robinson and Austen, 1987). Depleted sites of signal peptides that are recognized by signal peptidase have three conserved positions that are essential for cleavage but not for insertion in the membrane, these being the positions -1/-3/-6 (v. Heijne, 1984; Kuhn and Wickner, 1985). A small neutral amino acid is in position -1 and an uncharged (also larger) amino acid in position -3 and a helix-breaking residue (glycine or proline) is often situated in position -6. Rules for prediction of signal peptide cleavage sites were developed by v. Heijne (1983, 1986b). The important features of a signal peptide (polarity, secondary structure) could be summarized and functionally employed in a synthetic signal peptide which represents a derived consensus of known natural signal peptides. This peptide competes in vitro with numerous newly biosynthesized secreted proteins and prevents their take-up in microsomal vesicles and processing (Austen et al., 1984). By amino acid substitution, the natural hydrophobic core sequence for alkaline phosphatase from *E. coli* was mutated so that a sequence of nine consecutive leucine residues was now present. This mutant exhibited an increased secretion rate relative to wild type (Kendall et al., 1986). Kaiser et al. (1987) found that 20% of the statistical sequences are capable of assuming the function of a signal peptide as an export signal, in which the h region must comprise only a few hydrophobic residues, but only a few of them were accessible to processing by signal peptidase. These data together indicate that the specificity of signal peptide recognition is only low. Nevertheless, Bird et al. (1987) demonstrated an example in which a yeast signal peptide in the original state did not function as a secretion signal in a mammalian cell system, but only after mutation to higher hydrophobicity.

Analogous to the signal peptide, sequences were also found that cause insertion of the corresponding protein in the membrane and prevent secretion (stop transfer sequence). This applies to prokaryotes (Dalbey and Wickner, 1986; Kuhn et al., 1986) and eukaryotes (Mize et al., 1986; Munro and Pelham, 1987). In the mentioned examples, these stop transfer sequences are found on the C-terminus of the mature protein. By doubling of a signal peptide, depending on the spacing of the two sequences from each other, the second sequence could assume the function of a stop transfer sequence (Coleman et al., 1985). Szczesna-Skorupa et al. (1988) mutated a membrane-anchoring signal peptide by introduction of positive charges on the amino terminus to a secreting peptide. For the further transport path of secreted proteins, the participation of vesicles both between the rough endoplasmic reticulum and the Golgi apparatus, the next station of exocytosis (Lodish et al., 1987, 1988), and also further transport from there to the cell membrane as well as the merging of such vesicles with the cell membrane (Staehelin and

Chapman, 1987) were detected. In addition to the cell organelles, cytosol factors and ATP, as the energy-supplying compound, are necessary, as was demonstrated in an in vitro reconstitution experiment (Beckers et al., 1987). The distinction between plasma membrane proteins, secreted proteins and lysosomal proteins, which originally all possess a signal peptide, is attributed to a sorting toward different vesicle types in the last Golgi compartment, the trans-Golgi network (Griffiths and Simons, 1986). Rothman (1987) controversially discusses the different approaches of protein sorting and prefers the idea of selective retention of ER and Golgi proteins in these organelles by special retention signals. For bacteria, a varying hydrophobicity of the N-terminal segment of the signal peptide was related to final localization of the protein (Sjostrom et al., 1987).

A second aspect is retention of incorrectly folded or aggregated proteins in the ER. An example for such transient retention is the discovery of immunoglobulin heavy chain-binding proteins (BiP) in the lumen of the ER (Haas and Wabl, 1983), this interaction only being eliminated by correct aggregation of two light and two heavy chains (Bole et al., 1986). The C_{H1} domain of the heavy chain is necessary for interaction with BiP, but BiP is not necessary for assembling of light and heavy chains, and if this interaction does not occur, secretion of incompletely aggregated antibodies also occurs (Hendershot et al., 1987). A reason for this is that the target-determining sequences in the heavy chain are not always readily accessible. Assembling is also inhibited by incomplete glycosylation of the heavy chain in IgM. Dorner et al. (1987) demonstrated that the degree of interaction between BiP and the heavy chain depends on the degree of glycosylation; the lower the glycosylation of the heavy chain, the more stronger the binding to BiP, which can lead to a stable bond. The decisive factor is obviously the conformation and with it the accessibility of the glycosylation sites. The efficiency of secretion is regulated by these mechanisms.

Glycosylation is a key factor during protein sorting (Olden et al., 1982). This occurs in steps in the ER and Golgi apparatus (Ferro-Novick, 1985), in which different proteins impose different requirements on carbohydrate modification for their proper transport (Olden et al., 1978; Gibson et al., 1979). Guan et al. (1985) was able to demonstrate that by N-glycosylation an originally membrane-anchored protein could be altered so that it is efficiently transported to the cell surface. For this purpose, two consensus sequences (Asn-X-Ser/Thr) were introduced for N-glycosylation by site-directed mutagenesis. Wild type yeast cells synthesize two forms of invertase: a cytoplasmic form and a secreted glycosylated form with a signal peptide; on introduction to the ER lumen, the signal peptide is cleaved and 9-10 oligosaccharide chains are bound to asparagine. External chain oligosaccharides are linked to the N-bound inner carbohydrate chains during passage through the Golgi apparatus; active conformation cannot be

achieved when translocation occurs in the absence of glycosylation (Ferro-Novick, 1985). Glycosylation is obviously necessary to achieve a stable tertiary structure and for secretion.

Only a few studies have been conducted, especially for plant cells. Gubler et al. (1986) was able to demonstrate for the first time by immunogold labeling the participation of the Golgi apparatus in secretion of plant tissue. Signals were observed over the ER-lumen and associated with membranes of the rough ER and the Golgi apparatus. Similarly, through the mechanism of translocation through the ER membrane in plants, SRP itself was initially identified as a component of mammalian cells; in microsome membranes of wheat germ a component similar to SRP was discovered (Prehn et al., 1987), which, however, causes no translation stop. This could only be observed during addition of mammalian cell SRP (Duong et al., 1987). Only recently N. Campos (personal communication) reported the isolation of SRP from corn endosperm.

The best investigated secreted protein in plants is α -amylase (EC 3.2.1.1) from the aleurone layer of barley. Synthesis of this enzyme is stimulated by gibberellic acid up to a fraction of 50% total protein, in which case de novo synthesis occurs (Mozer, 1980; Bernal-Lugo et al., 1981; Higgins et al., 1982; Chandler et al., 1984; Huang et al., 1984). Several cDNA clones that belong to two different gene families have since been isolated (Rogers and Milliman, 1983; Chandler et al., 1984; Huang et al., 1984). In the present study clone E of Rogers and Milliman was used. The protein contains a signal peptide of 23 amino acids whose gene sequence was ligated in front of the structural genes of the proteins being investigated. Localization in the aleurone tissue has already long been investigated, but methods that permit adequate fine structure analysis were not always chosen. On the one hand, an association with aleurone grains, especially the outer membrane and peripheral cytoplasm (Jacobsen and Knox, 1973), was found by immunofluorescence and, on the other hand, signals in the core region and the entire cytoplasm (Jones and Chen, 1976). The participation of the ER (Locy and Kende, 1978; Jones and Jacobsen, 1982) was initially detected by biochemical methods, participation of both the ER and the Golgi apparatus finally by electron microscopy (Fernandez and Staehelin, 1985) and immunogold labeling (Gubler et al., 1986) – i.e., retention of a "conventional" secretion path. The liberation of α -amylase from aleurone cells by gibberellic acid-induced cell wall hydrolysis was recently demonstrated by immunogold labeling (Gubler et al., 1987). Processing and secretion of intact wheat α -amylase in yeast was achieved by Rothstein et al. (1984) by cloning of cDNA in a yeast expression vector.

Lysozyme, as a simple monomer and small protein from which a resistance effect can simultaneously be expected, was chosen for expression in tobacco. It is used for lysis of bacteria in pharmaceutical preparations and in in vitro laboratory methods. Lysozyme from the bacteriophage T4 (Tsugita, 1971; Grutter et al., 1983) is superior in strength of the effect to the chicken protein lysozyme ordinarily used because of its good availability – it is 250 times more

active than chicken protein lysozyme against *E. coli* cell walls. The T4 lysozyme is a basic protein of 18.7 kDa with an optimal pH of 7.2-7.4 for enzymatic activity. Although there are no homologies to chicken protein lysozyme in the primary structure, both proteins have a very similar conformation (Remington and Mathews, 1978; Mathews et al., 1981). At identical catalytic activity, similarities occur in the substrate binding mechanisms (Remington and Mathews, 1978). By introduction of a disulfide bridge, Perry and Wetzel (1984; Wetzel et al., 1988) could stabilize the protein against thermal inactivation. Purification of T4 lysozyme was first accomplished in 1968 (Tsugita et al., 1968).

During the infection cycle, the gene in *E. coli* is transcribed in early and late phase but the early mRNA is not translated because it forms a stable secondary structure that blocks the translation start by a "hairpin loop" (McPheeters et al., 1986). A genomic clone for the T4 lysozyme (clone E) was isolated a few years ago and sequenced (Owen et al., 1983) and this clone was used in the present study. Perry et al. (1985) achieved nontoxic expression of the same gene in *E. coli*. For this purpose, they cloned the gene with elimination of the hairpin-forming region behind the tac promoter and weak expression of the gene proved to be nontoxic. After induction with IPTG, up to 2% of the total protein formed as active T4 lysozyme but lysis only occurred after freezing and rethawing. This supports the assumption that expression of other T4 genes might be necessary for lysis of the host cell (Josslin, 1970; Tsugita, 1971). The T4 lysozyme overproduced in *E. coli* by Perry et al. and isolated from it was injected into rabbits to generate polyclonal antibodies and from this an affinity-purified IgG fraction was isolated.

A natural (although generally very weak) lysozyme activity has been found in a number of plants. This is exerted by proteins which primarily exhibit a chitinase activity. Such proteins are active against bacterial and fungal cell walls (Howard and Glazer, 1967; Glazer et al., 1969; Bernier, 1971; Boller et al., 1983; Pilet et al., 1983; Pilet and Bernasconi, 1984; Bernasconi et al., 1986; Schlumbaum et al., 1986). They belong to proteins whose expression is stimulated by stress conditions or attack of a pathogen (Bernier, 1971; Bernasconi et al., 1986; Schlumbaum et al., 1986) (see above, PR protein).

With integration of the gene for bacteriophage T4 lysozyme and its expression in tobacco, a contribution should be made to strengthening the natural defense mechanisms, especially a supplementation of endogenous lysozyme activity.

More intricately structured molecules are the antibodies (see Tijssen, 1985) consisting of two light and two heavy chains bound to each other by disulfide bridges. A distinction is made between several classes of heavy (50-70 kDa) and two classes of light chains (about 25 kDa). Additional stabilization occurs by noncovalent interactions. The individual chains are constructed from several domains, each of which comprises about 110 amino acid residues. They have significant sequence homologies and similarities in the three-dimensional structure between

the individual classes. The light chains include two domains, the heavy chains three to five. One distinguishes V, J, D and several C domains, the V region forming the active center and having the least sequence conservation. Hypervariable regions are arranged around the active center. The other domains are "joining," "diversity" and "constant." Because of the compact globular structure, which is formed primarily by the backbone of the constant region, antibodies acquire unusual stability. Even after chemical blocking of the previously cleaved S-S bridges, the chains reassociate the intact molecules, which demonstrates the intensity of the secondary interactions between adjacent domains. An exposed site is the "hinge" region, which is characterized by proline abundance, which leads to an unstable secondary structure, and cysteine residues, which bond the two heavy chains to each other. This position is relatively more sensitive to cleavage by proteases.

An anti-NP-IgM antibody described by Bothwell et al. (1981, 1982) was used as the model system (NP = (4-hydroxy-3-nitrophenyl)acetyl). The cDNAs of the gene for the light and heavy chain were made available by K. Rajewski (Cologne). Expression of both chains was already successful in *E. coli*, but assembling could only be found by in vitro reassociation from an extract (Boss et al., 1984). Cabilly et al. (1984) reached the same result with another antibody also in *E. coli*. The expression experiment of anti-NP antibody in yeast cells was successful (Wood et al., 1985). Synthesis, processing and secretion of the light and heavy chains, glycosylation of the heavy chain and the detection of functional antibodies has been demonstrated, but assembling has only occurred with low efficiency.

In a previous project, M. Stieger (1987) investigated expression of light and heavy chains as well as their assembling in *Nicotiana tabacum*. For this purpose, the sequence for the signal peptide was cleaved from the coding region for both chains and 30 amino acids removed from the carboxyl terminus of the heavy chain. Under control of a plant promoter, the presence of both chains could be demonstrated with weak signals. Stieger was not in a position to demonstrate assembling of a functional antibody in tobacco plants. Only in identically transformed *Acetabularia* has assembling already been detected.

Starting from these construct, fusion genes with the sequence for the signal peptide of α -amylase were produced from barley. Because of this, a secretion process in the plant should be introduced, as expected, which first allows us to expect glycosylation of the heavy chain and then possibly favors assembling, among other things merely by higher local concentrations of both chains. As already described, an immunoglobulin-heavy chain-binding protein (BiP) has been identified in the lumen of the ER in hybridoma cells whose interaction is only eliminated by aggregation of two light and two heavy chains (Haas and Wabl, 1983; Bole et al., 1986). The presence of a similar general mechanism for other proteins in plant ER would be conceivable. The use of monoclonal antibodies as a resistance strategy for plants has since been pursued by

another team (D. Baulcombe, personal communication). Fusions of the α -amylase-signal peptide gene with the coding regions for the light chain, heavy chain and lysozyme occurred as a function of the available clones in different ways. For the light chain of the antibody, an ATG codon for methionine as the translation start and a 6-bp BamHI linker was constructed by M. Stieger (1987) after removal of the sequence that codes for the signal peptide, before the first amino acid of the mature protein. The α -amylase signal peptide gene was cloned directly in the reading frame in front of it so that the amino acids glycine, serine and methionine were additionally included in the fusion. A direct fusion of the corresponding gene to the signal peptide gene was constructed both for the heavy chain and for the lysozyme. In the case of the heavy chain, the coding sequence begins with the first amino acid of the mature protein, in T4 lysozyme with methionine. Thus far no clear decision is possible whether the signal peptidase also requires conserved amino acids on the amino terminus of the mature protein, i.e., in the most unfavorable cases some amino acids of the pure α -amylase, for substrate recognition. In the present study, such construction was dispensed with and in one case (light chain), the incorporation of the inserted linker-coded amino acids was also carried out.

The system of transformation of plants with genetically altered agrobacteria, which was developed first, is now already intensive, but has not been fully investigated by far. An attempt is made by determining the natural interaction between *Agrobacterium* and plants to improve and expand genetic engineering applications. The Ti (tumor-inducing) plasmid of *Agrobacterium tumefaciens* was identified as a functional factor, whose essential parts are the T-region and the vir-region (Hoykaas and Schilperoort, 1985). The T-DNA includes the DNA fragment enclosed by two 25-bp repeats, which is transferred to plants and integrated in their genome (Stachel and Zambryski, 1986d). Excision of this fragment from the Ti plasmid is accomplished, as was worked out recently, with the cooperation of genetic products of the vir-region (see Stachel and Zambryski, 1986d), which are again induced by wound-dependent produced plant substances (Stachel et al., 1985; Janssens et al., 1986). The interaction between the bacterium and the plant is initiated by trans-acting factors that are coded in the bacterium by the chromosomal virulence loci *chv A* and *chv B* (Douglas et al., 1985). In this case, the active role is played by the bacterium, whereas the plant cell acts passively, and for virulence and attachment of the agrobacteria, bacterial expression of auxin (IAA) is necessary (Matthysse, 1986). It could be demonstrated that the auxin does not act on the plant cell, but on the bacterium. Matthysse deems the participation of bacterial surface proteins and lipopolysaccharides as probable. Several hundred agrobacteria are bound per plant cell, but only a small fraction of all interactions leads to a DNA transfer (Depicker et al., 1985). The authors conclude that a plant cell is only transformed by one or at most a few bacteria and several T-DNA copies in a plant cell clone originate from an original T-region copy. Chemotaxis of agrobacteria with the Ti plasmid

pTiB6S3 to the phenolic plant wound exudate acetosyringone was recently discovered (Ashby et al., 1987), and the Ti plasmid function is at least partially responsible for this. The optimal concentration of acetosyringone is, at 10^{-7} M, two orders of magnitude lower than that for induction of the vir operon described below, which ultimately leads to transfer of T-DNA. The vir-region of octopine-Ti plasmid pTiAch5 is the best investigated. Vir B, vir C, vir D and vir E are expressed in bacteria only after activation by plant cells, vir A is constitutive and noninducibly regulated, whereas vir G is constitutively expressed and at the same time plant-inducible (Stachel and Nester, 1986c). Engstrom et al. (1987) conducted a more detailed study of vir-coded proteins. Vir A codes for a 91-kDa protein (localized in the cytoplasmic membrane fraction), which presumably acts as a membrane sensor protein for the plant signal molecule (Melchers et al., 1987), whereas the vir G protein activated by it positively regulates vir transcription (Stachel and Zambryski, 1986a). The vir B and C loci are involved, in addition to *chv* A and B, in the attachment and penetration of agrobacteria (or the Ti plasmid) into the host plant, whereas vir E regulates transposition of T-DNA in the host genome and a transposase enzyme could be involved (Gardner and Knauf, 1986). The vir D locus is essential for T-DNA circularization and these gene products must possess recombinase or exonuclease activity (Porter et al., 1987; Veluthambi et al., 1987; Yamamoto et al., 1987). Vir C is attributed a relation with host specificity of different species of agrobacteria (Stachel and Nester, 1986c). Mostly acetosyringone and α -hydroxyacetosyringone were detected as inducing compounds liberated by wounded plants and these act as an inducing agent for the vir operon (Okker et al., 1984; Stachel et al., 1985, 1986a; Holten et al., 1986).

The T-DNA element is characterized by almost perfect 25-bp repeats (borders) flanking both sides. Nopaline plasmids possess a T-DNA region. Octopine plasmids, on the other hand, have two independent regions (T_L and T_R regions). The function of these repeats has largely been explained, but not the question of whether this involves an essential component of the transfer system. Peralta and Ream (1985) demonstrated that the borders function efficiently only in a (wild type) orientation, whereas Rubin (1986) established independence from orientation. The right border repeat is then more important than the left (Caplan et al., 1985) and it alone causes efficient T-DNA transfer. Deletion of the right border while maintaining the left one leads to a drastic reduction in effectiveness of T-DNA transfer (Horsch and Klee, 1986). Indications have recently accumulated that the border repeats are not essential components, but only serve to increase the efficiency of transfer and to delimit the transferred region (van Haaren et al., 1987; R. Hain, personal communication). The sequences adjacent to the repeats outside of T-DNA then play an important role (Wang et al., 1987b). Buchanan-Wollaston et al. (1987) discovered the transfer of plasmid DNA from agrobacteria into plants by the action of *ori* T and *mob* functions of the Ti plasmid with a large host region with the presence of T-DNA borders. The authors

assume that the vir genes of the Ti plasmid take over the function of the tra genes during interbacterial gene transfer. The sequences in the vicinity of the borders obviously also exert an effect and Peralta et al. (1986) were able to identify an enhancer-similar sequence (overdrive) that lies directly adjacent to the right border and stimulates the activity of T-DNA transmission. These results could be confirmed by van Haaren et al. (1987). After induction of the vir genes by acetosyringone, specific nicks occur as a result in the borders (Stachel et al., 1986b; Wang et al., 1987a). The complementary strand is newly synthesized between these nicks and the old strand liberated as a so-called "T-strand" (Stachel et al., 1986b) and another model explains the occurrence of the repeatedly detected T-DNA circles by recombination on the borders (Koukolikova-Nicola et al., 1985; Machida et al., 1986; Stachel et al., 1986b). A fragment of the vir region includes genes for four proteins which are necessary and sufficient for production of the T-DNA circles (Alt-Moerbe et al., 1986). The following steps of transfer require further explanation. The limited host region of some Ti plasmids is due to a defect of the cytokinin gene of T-DNA (Hoekema et al., 1984; Yanofsky et al., 1985). The observation that no acetosyringone could be detected and as a result no T-DNA circularization in a number of investigated monocotyledonous plants that normally cannot be attacked by agrobacteria appears to be important (Usami et al., 1987). In addition, the transformation rate in *Arabidopsis thaliana* was significantly increased by addition of acetosyringone (Sheikholeslam and Weeks, 1987). This compound is therefore assigned decisive importance in transformation experiments. The single monocotyledonous cell culture transformed by agrobacterium described thus far originates from *Asparagus officinalis* (Hernalsteens et al., 1984).

In addition to the opine genes, T-DNA also includes additional conserved genes whose expression lead to synthesis of plant hormones (Leemans et al., 1982; Morris, 1986). Genes 1 and 2 code for two enzymes of the auxin synthesis pathway (Schröder et al., 1984; Kemper et al., 1985; van Onckelen et al., 1986) and gene 4 for an enzyme of the cytokinin biosynthesis pathway. Auxins and cytokinins are not only produced in plants by these genes, but also in agrobacteria which secrete these substances, and they were also surprisingly detected in other microorganisms (see Weiler and Schroder, 1987). Transfer and expression of phytohormone genes from agrobacteria into plants lead to undifferentiated tumor growth, which is regulated by the concentration and ratio of both genes (Amasino and Miller, 1982; Akiyoshi et al., 1983).

The opine genes, according to which different types of Ti plasmids are classified (for example, octopine, nopaline), are localized in octopine plasmids in the T_R DNA and in plant tissue lead to synthesis of unusual amino acids that cannot be broken down by the plants themselves, but merely serve as nutrients for the agrobacteria (see Weiler and Schroder, 1987). The best known opine is nopaline, whose nopaline synthase gene was isolated very early (Depicker et al., 1982; Bevan et al., 1983). A double promoter (pT_R) which regulates the 1' and 2'

genes for enzymes of the mannopine biosynthesis pathway acquired greater significance recently (Ellis et al., 1984; Karcher et al., 1984; Winter et al., 1984). These two transcripts saw the greatest frequency of all T_R-DNA transcripts. Velten et al. (1984) isolated a 479-bp fragment from pTiAch5 that contains both complete promoters, and constructed plant selection expression vectors with it (Velten and Schell, 1986). Good activity of the promoter was found in calli. Since then, the promoter has found use for expression of CAT in carrot protoplasts (Boston et al., 1987), *Bacillus thuringiensis* endotoxin (Vaeck et al., 1987), a streptomycin resistance (Jones et al., 1987) and bacterial luciferase from *Vibrio harveyi*, a heterodimeric protein where both coding sequences are under the control of both promoter directions (Koncz et al., 1987). Weak activity in agrobacteria and *E. coli* was also described for the analogous 2' promoter from pTiB₆806 (Gelvin et al., 1981, 1985) and an activating element was identified by deletion analysis in the 5'-untranslated region, which can also make a rudimentary octopine synthase promoter functional again (DiRita and Gelvin, 1987). By analysis of luciferase expression in different parts of plants, a differentiated expression and a possible auxin dependence of double promoter activity was found (A. Szalay, C. Koncz, personal communication). This was a starting point for precise analysis of the transcription intensity in different plant tissues and testing that possible regulation and inducibility under sterile culture conditions. Tissue-specific expression of the T-DNA promoter was recently found for another promoter, namely that of gene 5 of T_L-DNA from pTiAch5 (Koncz and Schell, 1986). This promoter is active in the callus and stem and only slightly in fully developed leaves. On a medium with high auxin and low cytokinin, expression in leaf fragments could be reproduced and it appears to be regulated by plant growth factors. Analysis of the expression conditions of the T_R double promoter was necessary in order to be able to establish in which tissue and under which conditions the introduced proteins can be expected. A tissue-specific regulation and inducibility was found by wounding and plant hormones. The nopaline synthase promoter (pNOS) is generally viewed as a constitutive, relatively weak promoter (Depicker et al., 1982; Bevan et al., 1983; Koncz et al., 1983; Sanders et al., 1987). This promoter also has an activating element (Ebert et al., 1987). Since the gene for the heavy chain of the antibody in the existing transgenic plants is under the control of the NOS promoter, an attempt was made to transfer the induction conditions for the double promoter to the NOS promoter. Both promoters are a component of T-DNA and also control synthesis of opines and therefore the result of a logical line of thought developed to also investigate pNOS for possible regulation and inducibility, despite all previous publications. In fact, the same expression pattern in principle was found, but with much lower factors but at roughly the same base level. The previous classification as a simple constitutive promoter is therefore incorrect and became untenable.

The plants generated during a transformation (even those from a callus) consequently exhibit very high variability in expression of the introduced gene (see Budar et al., 1987; Sanders et al., 1987). It is therefore essential to test all transformants for the expression rate.

So-called disarmed Ti plasmids were constructed for genetic engineering use of the Ti plasmid system. They do not generate tumorous growth and permit regeneration of intact plants. The already constructed plasmid pGV 3850 (Zambryski et al., 1983) is also still in use today. The recombinant vector plasmid that contains the promoter-gene-pA sequence cassette is integrated by homologous recombination. More recent constructions function according to the principle of binary vectors, where the borders are localized with the T-DNA on a small plasmid and the vir region in the trans position on a second plasmid (Bevan, 1984; Hoekema et al., 1985; Koncz and Schell, 1986; Koncz et al., 1987).

A rapid method for detection of gene products can be transient expression of a plasmid in tobacco protoplasts. This method was first worked out recently (Fromm et al., 1985; Prols, 1986; Fromm and Walbot, 1987) and is derived from the second established gene transfer method for transformants (in addition to the agrobacteria system), the direct gene transfer with naked DNA (Krens et al., 1982; Paszkowski et al., 1984; Hain et al., 1985; Shillito et al., 1985; Negutiu et al., 1987). During transient expression, the genetic information is read directly from the free plasmid and integration into the genome is unnecessary. After a short time, however, the DNA breaks down. In such a system, expression of NPT II with the T_R double promoter of comparable intensity has already been measured with the cauliflower mosaic virus 35S promoter (M. Prols and R. Topfer, personal communication) and the calcium nitrate-PEG transformation methods modified according to Hein et al. (1983) were used. Transient expression of a plasmid-coded protein by Western blotting, however, has still not been described in tobacco protoplasts.

Another rapid experiment is expression of such plasmids in *Acetabularia mediterranea*. The plasmid DNA is injected into isolated nuclei, which are then reimplanted in *Acetabularia* without nuclei (Neuhaus et al., 1983, 1984, 1986; Langridge et al., 1985; Berger et al., 1987; Schweiger and Neuhaus, 1987). Detection of the heterologous protein occurs by immunofluorescence using appropriate control. Such an experiment has been successfully conducted for the T_R double promoter-T4 lysozyme gene fusion by G. Neuhaus (Ladenburg).

Identification of the sought proteins and detection of cleavage of the signal peptide can occur by analysis of the molecular weight and immunochemical detection by Western blotting (Towbin et al., 1979; Burnette, 1981). This method has found increasingly successful use in recent years in molecular biology laboratories. A further development is the recently introduced semidry electroblotting (Kyhse-Andersen, 1984), combining simplicity and rapidity of the method. A modern and reliable Western blotting method with biotin-streptavidin-supported immunochemical detection by alkaline phosphatase staining was developed together with it. The

use of radioactive labeling was also applied to DNA analysis by Southern blotting and a similar protocol worked out. For enrichment of B1-8 antibodies from plant tissue, the most promising method has proved to be the use of monoclonal anti-B1-8 antibodies (A. Radbruch and M. Reth, personal communication; Reth, 1981). An affinity chromatographic method was worked out for it and also similarly for T4 lysozyme.

Final demonstration of the localization of these proteins in plant tissue can only occur by electron microscopic methods. Since only relatively limited amounts of protein are to be expected in comparison with previous analyses of endogenous protein, the most sensitive method, immunogold labeling, was resorted to. Analyses have already been conducted on plant material with such a method (Sossountzov et al., 1986; Gubler et al., 1987a, 1987b).

In the present study, gene constructs were initially prepared, which, on the one hand, have the genes for the light chain and the heavy chain of an anti-NP monoclonal antibody (B1-8) under the control of the T_R double or NOS promoter and, on the other hand, for T4 lysozyme under the control of the T_R double promoter. All genes are fused with the gene for the signal peptide of α -amylase from barley in order to achieve secretion of the gene products. These expression vectors were integrated by homologous recombination into the Ti plasmid pGV 3850 of agrobacteria. By the leaf disk test, *Nicotiana tabacum* was transformed individually with these constructions and transgenic plants regenerated. By Southern blot and the NPT tests (NPT II gene under control of the second position of the T_R double promoter), the transformation was checked and confirmed. A promoter activity analysis for p T_R and pNOS was conducted on transgenic plant material to determine in which tissue and under which conditions expression is to be expected, according to which it turned out that the double promoter does not operate constitutively. Based on these results and the thus-far isolated transformed plants, a naked DNA transfer was carried out in *Nicotiana tabacum* SR1 for the T4 lysozyme construct in cooperation with R. Hain (Monheim). Parallel with generation of the genetically altered plants, a polyclonal anti-T4 lysozyme antibody was prepared and characterized from rabbits. Modern, highly sensitive, nonradioactive detection methods based on the biotin-streptavidin system including semidry electroblottings for Western and Southern blotting were worked out for analysis of the plants. For enrichment of the sought proteins from the crude extract, an affinity chromatographic method proved to be the method of choice. Monoclonal Sepharose-coupled anti-B1-8 antibodies and a hapten-Sepharose gel were used for B1-8 (furnished by A. Radbruch and M. Reth). An attempt was also made to prepare a homologous system of transient expression in tobacco mesophyll protoplasts. The sought proteins were enriched by the aforementioned method and analyzed in this form or partly from the crude extract by Western blotting. In order to localize the proteins in the cell and demonstrate secretion, a joint project was conducted with S. Hippe (Aachen) in the area of electron microscopic immunogold labelings. For this purpose appropriate

and modern preparation techniques were worked out by S. Hippe in order to obtain the structure of the cells from the different plant tissues and at the same time prevent possible loss of intercellularly localized proteins.

II. Material and methods

2.1 Bacterial strains

E. coli strains

- DH I (F⁻, rec A1, end A1, gyr A96, thi-I, hsd R17, sup E44, rel A1, λ)
 Low (1968), Hanahan (1983)
 In this strain (rec A⁻) the danger of recombinations in the cloned DNA is sharply reduced.
- BMH 71-18 Δ (lac-pro aB), thi, sup E, (F⁻, pro AB, lac I^a, Z Δ M15)
 Yanish-Perron et al. (1985)
 In this strain blue-white selection of transformed bacteria is possible with an insert in the lac Z gene by means of IPTG/X-Gal.
- GJ 23 (R 64 rdl1, pGJ 28, Sm^R, Tc^R, Km^R)
 Van Haute et al. (1983)
 This strain contains the helper plasmids for conjugation of *E. coli* with agrobacteria.

Agrobacteria strains

- C 58 Cl (pGV 3850, Rif^R, Cb^R), Zambryski et al. (1983)
 GV 2215 (Rif^R), Leemans et al. (1982)

2.2 Plant material

Nicotiana tabacum Wisconsin W38 (W38) is cultured sterile from seeds and reproduced as a sprout culture on MS medium (Murashige and Skoog, 1962) for a full day with 16 h light and 8 h darkness at 24-26°C.

Nicotiana tabacum var. *petit Havanna* SR1 (Maliga et al., 1973) is used for the protoplast studies and cultivated in the same manner.

2.3 Chemicals

General chemicals and auxiliaries were obtained from the following companies: Aldrich, Baker, Bio-Rad, Boehringer, Calbiochem, Difco, Fluka, Hoechst, Merck, Seakem, Serva, Sigma, Whatman, Schleicher & Schuell.

DNA-modifying enzymes, DNAs and linkers from: Boehringer, BRL, Pharmacia, New England Biolabs.

Radiochemicals from Amersham.

X-ray films: Kodak X-Omat S and X-Omat AR.

2.4 Cooperation

Acetabularia mediterranea: Dr. Gunter Neuhaus, Ladenburg (culture, microinjection, immunological detection).

Immunogold labeling was conducted in cooperation with Dr. Sigrun Hippe (Aachen). During this work, a high pressure-low temperature sample preparation was conducted in the laboratory of Prof. Hans Moor (Zurich).

T_R promoter analysis: Dr. Robert Masterson (Cologne).

Naked DNA transfer for stable transformation with the T4 lysozyme construct:
Dr. Rudiger Hain (Monheim).

Expression of one-chain antibodies in vitro in plants: Prof. Georges Kohler and Dr. Antonio Iglesias (Freiburg).

Obtained materials:

The following DNA clones were available:

Clone E (α -amylase from barley) from J. C. Rogers (Rogers and Milliman, 1983).

T4 lysozyme (λ 1358-b538-imm 434::T4, E III e' sus 57) from G. R. Smith (Owen et al., 1983).

pAB λ 1-15 (light chain of B1-8) and pAB μ -11 (heavy chain of B1-8) as well as B1-8 protein and the employed antibodies and affinity gels for it from K. Rajewski (Cologne) (Bothwell et al., 1981, 1982).

Plasmids pLGV 2385-Lc, pHC 1-3 and pHC 3-19 from M. Steiger (Cologne).

T4 lysozyme protein from R. Wetzel (Genentech, San Francisco) (Perry et al., 1985).

2.5 Microbiological methods

2.5.1 Culturing of *E. coli*

E. coli bacterial strains are cultured in LB medium (Miller, 1972):

10 g tryptone

5 g yeast extract

5 g NaCl

H₂O to 1 L, pH 7.5

The medium is mixed with the required antibiotic. Employed concentrations:

Amp 100 mg/L

Sm 12.5 mg/L

Sp 50 mg/L

Incubation at 37°C in a blocking culture.

2.5.2 Culturing of agrobacteria

The agrobacteria strains were cultured in YEB medium.

5 g beef extract
1 g yeast extract
1 g peptone
5 g sucrose
2 mM MgCl₂
H₂O to 1 L, pH 7.4

or in isolated agrobacteria clones also in LB medium (Miller, 1972) (see 2.5.1).

The media are mixed with the required antibiotics. Employed concentrations:

Cb 100 mg/L
Rif 100 mg/L
Sm 300 mg/L
Sp 100 mg/L
Km 25 mg/L

Incubation at 28°C.

2.5.3 Conjugation between *E. coli* and *Agrobacterium tumefaciens*

The mobilization system of the helper plasmids R64drd11 and pGJ28 is described by Van Haute et al. (1983). These two helper plasmids are transferred by conjugation between two *E. coli* strains from the strain GJ 23 in DHI, which contains the vector being mobilized. In the next step, all three plasmids are transformed from these transconjugates in a conjugation between *E. coli* and *Agrobacterium tumefaciens* in the strain C 58 C1 (pGV 3850) (Zambryski et al., 1983). Both conjugation steps are conducted simultaneously in one procedure.

For this purpose, 5-mL overnight cultures were cultured from the bacterial strain (DH I with pAP 2034 derivative and GJ 23 in LB at 37°C, C 58 C1 (pGV 3850) in YEB at 28°C). 20 µL of the cultures of both *E. coli* strains are added dropwise in the center of a YEB plate and incubated for 1 to 2 h at 37°C, the spot drying in this time. 20 µL of the agrobacteria culture are then added dropwise to the center of the spot and incubated overnight at 28°C, suspended in 400 µL YEB medium and 200 µL of it placed on a YEB plate with Rif (100 mg/L), Cb (100 mg/L), Sm (300 mg/L) and Sp (100 mg/L) and incubated for 2 to 3 days at 28°C.

2.6 Plant tissue culture and transformation methods

2.6.1 Culturing conditions

Sterilization of tobacco seeds with sodium hypochlorite:

The seeds are placed for 2 min in 70% ethanol and 10 min in SDS sodium hypochlorite solution (2% (w/v) SDS + 17.5% (w/v) sodium hypochlorite in sterile water) and washed 2×5 min. The sterilized seeds are laid out on L + S medium (Linsmeier and Skoog, 1965) and germinated and proliferated in a sterile culture on M + S medium (Murashige and Skoog, 1962) with a 16-h day at 24-26°C.

For transformed material, L + S medium (Linsmeier and Skoog, 1965) with 3% sucrose for the callus culture and sprout induction and M + S medium (Murashige and Skoog, 1962) with 1% sucrose for plant regeneration and culture are used.

Composition	L + S	M + S
NH ₄ NO ₃	1650 mg/L	
KNO ₃	1900 mg/L	
CaCl ₂ × 2H ₂ O	440 mg/L	
MgSO ₄ × 2H ₂ O	370 mg/L	
KH ₂ PO ₄	170 mg/L	
Na ₂ EDTA	37.3 mg/L	
FeSO ₄ × 7H ₂ O	27.8 mg/L	
H ₃ BO ₃	6.2 mg/L	
MnSO ₄ × H ₂ O	16.9 mg/L	
ZnSO ₄ × 7H ₂ O	8.6 mg/L	
KI	0.83 mg/L	
Na ₂ MoO ₄ × 2H ₂ O	0.25 mg/L	
CuSO ₄ × 5H ₂ O	0.025 mg/L	
CoCl ₂ × 6H ₂ O	0.025 mg/L	
myo-inositol	100 mg/L	
Thiamine hydrochloride	0.4 mg/L	0.1 mg/L
Glycine		2 mg/L
Nicotinic acid		0.5 mg/L
Pyridoxine hydrochloride		0.5 mg/L
Sucrose	30 g/L	10 g/L
	pH 5.8	

2.6.2 Leaf disk test for regeneration of transformed plants (Horsch et al., 1985)

Method for direct generation of sprouts (modified according to Hain et al., 1985)

Younger leaves of *Nicotiana tabacum* W38 are cut into small pieces and introduced to petri dishes filled with liquid L + S medium (Linsmeier and Skoog, 1965). For this purpose

100 μ L of an overnight culture of agrobacteria are pipetted and incubated for 2 days. This is then washed three times in L + S medium with 0.5 mg/L BAP, 0.1 mg/L NAA and 500 mg/L cefotaxime (Claforan, Hoechst) and the leaf pieces placed on L + S medium with 0.5 mg/L BAP, 0.1 mg/L NAA, 500 mg/L Claforan (3% sucrose and 0.8% agar). After a week the leaf pieces are transferred to fresh medium that additionally contains 50 mg/L kanamycin for selection (the vector pAP 2034 contains a gene for kanamycin resistance (NPT II)). The direct formation of sprouts is induced by the employed hormone combination without forming larger calli. The formed sprouts are cut off and placed for further selection on M + S medium (Murashige and Skoog, 1962) with 2% sucrose, 100 mg/L kanamycin, 500 mg/L Claforan. Untransformed sprouts wither. Surviving sprouts are tested for the presence of nopaline (see 2.9.1), if the nopaline gene is present in the Ti plasmid. Nopaline-positive sprouts are rooted and further cultivated on the M + S medium with 500 mg/L Claforan. Surviving but nopaline-negative plants are also further treated. When the sprouts are large enough, they are tested for neomycin phosphotransferase II activity (NPT test/see 2.9.2).

Coinfection method (slightly modified according to Kaulen, 1986)

The vector pGV 3850 can be used together with the sprout-inducing agrobacteria mutant GV 2215 (Leemans et al., 1982). For this purpose, the leaf pieces are mixed with pGV 3850/GV 2215 in a 5:1 ratio and incubated in a liquid L + S medium. After 2 days, they are washed three times in liquid L + S medium and the leaf pieces are placed on L + S medium with 3% sucrose, 0.8% agar, 500 mg/L Claforan and 50 mg/L kanamycin. Two to three weeks after coinfection, tumors are formed which form sprouts in the medium in the absence of exogenous hormones. Initially, the calli are selected on 50 mg/L kanamycin, later the sprouts on 100 mg/L kanamycin. Further treatment of the sprouts occurs as described above.

With acetosyringone induction

The overnight culture of agrobacteria is cultivated under antibiotic selection in LB medium (pH 5.6) with 20 μ M acetosyringone (stock solution: 10 mM in DMSO). Infection of the tobacco leaves occurs with 100 μ L of this culture in L + S medium (pH 5.6) with 20 μ M acetosyringone for two days. It is then further treated according to the first method. Selection on 50 mg/L kanamycin begins after one week.

2.6.3 Transient expression in tobacco mesophyll protoplasts

Transformation of *Nicotiana tabacum* SRI protoplasts with naked DNA occurs slightly modified according to Prols (1986) with the calcium nitrate fusion method.

Younger leaves are cut to medium-sized pieces and digested overnight in the dark at about 24°C in enzyme solution (1% cellulase Onozuka R10 + 0.5% macerozyme in K₃ medium) (>0.4M sucrose, 600 mOs)). On the next morning, they are carefully shaken and digested for another 1 to 2 h. The solution is then passed through a 250-μm and a 100-μm sieve and then floated two times with K₃ medium and pelletized once with sea water (600 mOs). The protoplasts are set at a density of 5×10^6 cells/mL.

200 μL (1×10^6 protoplasts) of the suspension are transformed to a petri dish. For this purpose, 4 x 50 μL drops of PEG solution are placed around the protoplast drops and the DNA solution furnished to the protoplasts; with a larger volume of DNA solution, the amount of PEG solution is increased accordingly. This is then carefully mixed with the protoplasts and incubated for 20 min. 5 mL of the wash solution are then added dropwise, the entire liquid filled into a centrifuge tube and allowed to stand for 10 min. The protoplasts are centrifuged off, taken up in 3 mL of K₃ medium + 1 mg/L NAA + 0.2 mg/L kinetin and filled into a small petri dish. After a few hours, the protoplasts are separated from each other and cultivated in the dark for three days at 24°C. They are then centrifuged with at least one volume of a 0.24M CaCl₂ solution and further processed as described for the different analysis methods.

K₃ medium (Kao et al., 1974; Nagy and Maliga, 1976):

NaH ₂ PO ₄ × H ₂ O	150 mg/L	MnSO ₄ × H ₂ O	10 mg/L
CaHPO ₄	50 mg/L	H ₃ BO ₃	3 mg/L
CaCl ₂ × 2H ₂ O	900 mg/L	ZnSO ₄ × H ₂ O	2 mg/L
KNO ₃	2500 mg/L	Na ₂ MoO ₄ × 2H ₂ O	0.25 mg/L
NH ₄ NO ₃	250 mg/L	CuSO ₄	0.025 mg/L
(NH ₄) ₂ SO ₄	134 mg/L	CoCl ₂ × 6H ₂ O	0.025 mg/L
MgSO ₄ × 7H ₂ O	250 mg/L		
FeSO ₄ × 7H ₂ O	27.8 mg/L		
Na ₂ EDTA	37.3 mg/L		
myo-inositol		100 mg/L	
Nicotinic acid		1 mg/L	
Pyridoxine hydrochloride		1 mg/L	
Thiamine hydrochloride		10 mg/L	
Xylose		250 mg/L	
>0.4M sucrose, 600 mOs, pH 5.6			

PEG solution:

0.1M Ca(NO₃)₂
0.45M mannitol
25% PEG 6000

0.1M HEPES, pH 9.0

Wash solution:

0.275M $\text{Ca}(\text{NO}_3)_2$, pH 6.0

2.6.4 Stable transformation of tobacco mesophyll protoplasts

Transformation of SRI leaf protoplasts occurs as described under 2.6.3.

They are cultivated initially for 3 days in the dark, then in the light. A week after transformation, the regenerating protoplasts are embedded in agarose (1:1 (vol) with 1.2% SeaPlaque low-melting agarose in K_3 medium ($>0.4\text{M}$ sucrose, 600 mOs) with 1 mg/L NAA and 0.2 mg/L kinetin). After solidification of the agar block, this is divided into quarters and transferred to a large petri dish. The agar disks are placed in K_3 medium with 0.4M sucrose with hormones so that they are rinsed by the liquid medium. After 1 week the sugar content of the liquid medium is reduced by 0.05M sucrose until a value of 0.15M sucrose is reached. A week after embedding, selection for kanamycin begins (50 mg/L). After about 6 to 8 weeks, small calli should be visible, which are then removed from the agarose and further cultivated as described under 2.6.2.

2.6.5 Induction of pT_R and pNOS in plant tissue

Standard method for sterile culture:

Leaves: The leaves are cut off, carefully cut into small pieces with a scalpel and placed with the top down on solid M + S medium with Claforan and 10^{-5}M (2 mg/L) 2,4-D, as well as 10^{-5}M (2 mg/L) kinetin (1:1 hormone) and incubated for 5 or 6 days.

Stems: The stems are divided into roughly 1-cm-long segments, these are cut lengthwise and placed with the cutting surface down on solid M + S medium with Claforan without additives and incubated for 5 or 6 days.

Standard method for greenhouse plants:

Leaves: The leaf surface is roughened, for example, with a toothbrush and damaged in so doing, whereupon a solution of the aforementioned 1:1 hormone is sprayed on. On the two subsequent days the damaged leaves are sprayed with water. After 2 to 5 days, the induced leaves are cut off.

Test method for induction analysis:

Leaves and stem pieces are investigated in sterile culture according to the aforementioned methods with varying times (5 min to 6 days) and different additives (none, different hormone concentrations, different hormone concentrations and ratios, acetosyringone). Intact leaves and leaf pieces are used in a liquid and solid medium. The optimal concentrations are described above.

2.7 Methods of recombinant DNA technology

2.7.1 Isolation of plasmid DNA from *E. coli*

Cloned plasmid DNA from *E. coli* bacteria is modified according to Birnboim and Doly (1979) and isolated.

— Miniprep: preparation:

The bacteria are cultured in a 5 mL overnight culture (LB medium) at 37°C under antibiotic selection. 1.5 mL of this are filled into an Eppendorf vessel and centrifuged. The pellet is resuspended in 100 µL of solution I, mixed with 10 µL of solution II and incubated for 10 min at room temperature. 200 µL of solution III are then added, mixed on the vortex and left to stand on ice for 5 min, then precipitated with 150 µL of cold solution IV, left to stand for 5 min on ice and then centrifuged. The supernatant is treated once with the same volume of phenol/chloroform/isoamyl alcohol 25:24:1 and precipitated with 0.1 vol of 3M NaAc pH 5.0 and 2 vol of ethanol. After centrifuging, it is washed with 70% ethanol and dried, and the pellet is taken up in 30 µL of TE. About 5 µL of this are used per digestion.

Solution I: 10% (w/v) glucose
10 mM EDTA
25 mM Tris-HCl, pH 8.0

Solution II: 50 mg/L lysozyme in solution I

Solution III: 0.2M NaOH

1% (w/v) SDS

Solution IV: 3M NaAc, pH 4.8

— Maxiprep: preparation:

The bacteria are cultured in a shaking culture for about 20 h at 37°C in LB medium under antibiotic selection. The bacteria are harvested by centrifuging for 20 min at 3500 G, resuspended in 10 mL of solution I/L of culture and lysed with 3 mL of solution II/L of culture. This is incubated at room temperature for 30 to 60 min. After addition of 3 mL of 0.5M EDTA/L of culture, it is allowed to stand for 10 min at room temperature. After addition of 40 mL of solution III/L of culture, it is carefully mixed and incubated at room temperature for 15 to 30 min. It is centrifuged in polycarbonate tubes for 1 h at 65,000 G and 10°C. The clear supernatant is mixed with 0.97 g CsCl/mL and 0.083 mL EtBr (5 mg/mL)/mL and the plasmid DNA purified over two gradients. After dialysis versus water overnight, it is precipitated with alcohol, extracted twice with phenol/chloroform/isoamyl alcohol 25:24:1, precipitated with ethanol, washed with 70% ethanol, dried and dissolved in TE buffer. The concentration is determined via A_{260} . $A_{260} = 1$ corresponds to 50 µg DNA/mL.

Solution I: 25% (w/v) sucrose

50 mM Tris-HCl, pH 7.5

2.5 mM EDTA

Solution II: 10 mg lysozyme/mL in solution I

Solution III: 0.4% (w/v) SDS

50 mM Tris-HCl, pH 7.5

62.5 mM EDTA

2.7.2 Isolation of total DNA from agrobacteria

The agrobacteria total DNA is isolated from 5 mL of a two-day culture (28°C) in YEB or LB medium. The pellet is washed with 1.5 mL of 5M NaCl and resuspended in 300 µL of TE. 100 µL of a 5% (w/v) sarcosyl solution in TE and 100 µL of a predigested proteinase K solution (5 mg/mL) are added and digested for 2 h at 37°C. The DNA is then sheared by six- to eight-fold drawing through a 21 g cannula and extracted twice with phenol/chloroform/isoamyl alcohol 25:24:1 and washed with 70% ethanol. After drying, the DNA is dissolved in 100 µL TE and 10 µL of this is used for restriction digestion. The DNA is precipitated and stored under ethanol at 4°C.

2.7.3 Cloning methods

The fundamental methods are essentially performed according to Maniatis et al. (1982).

Recombinant plasmids are transformed in the competent cells (modified according to Maniatis et al., 1982, RbCl₂ method), pBR 322 and pAP 2034 derivatives in DH I or 71-18, pUC 9 derivative in 71-18. Transformation occurs in slightly modified fashion by a temperature shock (42°C) according to Dagert and Ehrlich (1979). Bacteria with recombinant plasmids are selected for antibiotic resistance. For identification of the sought clone, plasmid DNA is isolated with a minipreparation (see 2.7.1), digested with appropriate restriction enzymes and separated in agarose gels with 1× TBE buffer (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA). λ-DNA digested with Pst I generally serves as the size marker.

Employed plasmids:

pBR 322 (Bolivar et al., 1977)

pUC 9 (Vieira and Messing, 1982)

pAP 2034 (Velten and Schell, 1986)

Restriction digestions of DNA are conducted according to the information of the enzyme producing companies. The isolation of DNA fragments from agarose gels occurred by electroelution in a dialysis loop in 0.5× TBE. The DNA is purified in a DEAE column (DE 52, Schleicher & Schuell) and eluted with 1M NaCl. The eluate is eluted with phenol/chloroform/isoamyl alcohol 25:24:1 and precipitated with ethanol.

Ligations occur overnight at 15°C with 2 units T4 DNA ligase in 20 mM Tris-HCl, pH 7.6, 10 mM MgCl₂·10 mM DTT, 0.6 mM ATP in final volumes of 20 µL (about 1 µg DNA).

Fill in reactions with about 2 µg DNA are conducted with 10 mM dNTPs and 5 units DNA polymerase I Klenow fragment in 50 mM Tris-HCl, pH 7.2, 10 mM MgSO₄·0.1 mM DTT and 50 µg/mL BSA in final volumes of 75 µL for 30 min at room temperature.

Commercial linkers are acquired either in phosphorylated form or phosphorylated before ligation with T4 polynucleotide kinase. For this purpose, 2 µg of linker are incubated with 2 units of enzyme in 70 mM Tris-HCl, pH 7.6, 10 mM MgCl₂·5 mM DTT, 0.5 mM ATP for 45 min at 37°C in a final volume of 20 µL.

For linker ligation, 2 µg of filled in DNA and 1.5 µg phosphorylated linkers with 2 units T4 DNA ligase and 1 unit T4 RNA ligase are cultured overnight at 15°C. This is then extracted with phenol/chloroform/isoamyl alcohol 25:24:1, precipitated with ethanol and then "opened" with the corresponding restriction enzyme of the linker. After digestion, it is extracted with phenol/chloroform/isoamyl alcohol 25:24:1 and selectively precipitated with isopropanol. Only fragments >1 kB are precipitated with 0.55 vol isopropanol, with 0.6 vol from about 1 kB and larger unligated linkers remaining in solution. For this purpose, 2.5 µL of 0.5M EDTA/100 µL of reaction solution are added, mixed with isopropanol and allowed to stand for 30 min at room temperature. This is then centrifuged and washed with 70% ethanol.

Synthetic oligonucleotides were mechanically produced according to the phosphoramidite method at the Institute of Genetics at Cologne University. The work-up was as follows: The reaction solution after cleavage of the protecting groups is divided into 100-µL aliquots and ammonia evaporated overnight in an exhaust, then lyophilized in the Speed-vac and taken up again in 10 µL water. Purification occurs over a 20%-8M urea-polyacrylamide gel. The probes are then denatured for 2 min at 95°C, cooled on ice and placed on formamide-sampled buffer (80% formamide, 50 mM Tris-HCl, pH 8.3, 0.1M EDTA, 0.03% (w/v) bromophenol blue) and 1× TBE. The gel runs overnight with 450V. It is stained with 10 µL ethidium bromide (5 mg/L)/100 mL for 45 min in 1× TBE. The strongest and generally also the uppermost band is cut out on the transilluminator and eluted overnight in the roller with 0.5M NH₄Ac, 1 mM EDTA at 37°C. To eliminate gel particles, it is filtered through sterile glass wool precipitated for one day at -70°C with ethanol, then washed twice with 70% ethanol, dried and dissolved in H₂O. For ligation, the two complementary oligonucleotides are phosphorylated together in a solution with T4 polynucleotide kinase and ligated in the same charge after addition of DNA with T4 DNA ligase.

2.7.4 Southern blot

With nitrocellulose (conventional blot):

The DNA being analyzed is digested with the appropriate restriction enzyme and separated on an agarose gel in 1× TBE (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA). For depurination, it is washed twice for 15 min with 0.25M HCl, then denatured for 45 min with 1.5M NaCl/0.5M NaOH and neutralized for 45 min in 1M Tris-HCl, pH 7.0/1.5M NaCl. Transfer of DNA to nitrocellulose (Schleicher & Schuell BA 85, 0.45 μ m) occurs according to Southern (1975) and Maniatis (1982) by capillary blotting in 20× SSC 6 h to overnight. After blotting, the filter is briefly rinsed in 6× SSC, dried in air and baked for 2 h in a vacuum furnace at 80°C.

20× SSC: 0.3M sodium citrate, 3M sodium chloride, pH 7.0.

With Immobilon PVDF (alkaline blot):

In this case, the agarose gel after the run is also depurinated in large fragments (1× 15 min 0.25M HCl) and then denatured for 30 min in 0.4M NaOH. Transfer occurs in the same buffer. The Immobilon membrane (Millipore) is moistened briefly in methanol and the methanol substituted by immersing in water (about 2 min), and then the filter is equilibrated in the blotting buffer. The blot is constructed as usual and allowed to stand overnight. After blotting, the filter must be washed for 2 min in 2× SSC so that a neutral pH value is reached, since otherwise brown coloration of the filter occurs during baking. For fixation of the DNA, there are two possibilities to choose from: either baking for 2 h in a vacuum furnace at 80°C as in nitrocellulose, after the filter was dried at room temperature, or placing the moist filter for 4 min on a transilluminator.

2.7.5 Hybridization of filter-bound plasmid and agrobacteria DNA with biotin-labeled samples

During use of biotinylated probes, phenol must not be used for extraction, since they preferably migrate into the phenol phase!

— Biotinylation of DNA fragments by nick translation:

In an Eppendorf vessel 5 μ L Sol A4 (0.2 mM dATP, 0.2 mM dCTP, 0.2 mM dGTP, 500 mM Tris-HCl, pH 7.8, 50 mM MgCl₂, 100 mM 2-mercaptoethanol), 2.5 μ L 0.4 mM biotin 11-dUTP (BRL) and DNA are combined with 5 μ L nick translation-grade DNA polymerase I (BRL) in a total volume of 50 μ L. This is incubated for 60 min at 15°C, then the reaction terminated with 3 μ L 0.5M EDTA and 1.25 μ L 5% (w/v) SDS. Precipitation is done twice with ethanol to eliminate free nucleotides.

— Biotinylation of DNA fragments by oligolabeling (modified according to Feinberg and Vogelstein, 1983/1984):

The random primed DNA labeling kit from Boehringer is used. 25 ng to 1 µg DNA are denatured for 10 min at 95°C and cooled on ice. 1 µL of the dATP, dCTP and dGTP solutions (5 mM), 1.4 µL of a 1:0.6 dTTP/bio-11-dUTP mixture (5 µL 0.5 mM dTTP + 3.75 µL 0.4 mM bio-11-dUTP), 2 µL of reaction mixture and 1 µL of Klenow enzyme are added and allowed to react in a total volume of 20 µL for 2 h at 37°C. The reaction is stopped by adding 1 µL of 0.5M EDTA and 10 min of heating at 65°C.

— Hybridization with biotinylated DNA probes:

Nitrocellulose filters are moistened in water, Immobilon filters are wetted in methanol and substituted in water and then prehybridized for 2 h at 42°C in hybridization buffer (5 mL/100 cm of filter) with heat-denatured (10 min at 95°C) sonicated calf thymus DNA (300 µg/mL of buffer). Overnight hybridization is done in fresh hybridization buffer (2 mL/100 cm filter) with the heat-denatured (10 min at 95°C) DNA probe and sonicated calf thymus DNA (100 µg/mL of buffer). The washings (stringent) are as follows: 2 x 3 min in 2× SSC/0.1% SDS at room temperature, 2 x 3 min in 0.2× SSC/0.1% SDS at room temperature and 2 x 15 min in 0.16× SSC/0.1% SDS at 50°C. After brief rocking in 2× SSC at room temperature, it is either detected or dried in air.

Hybridization buffer:

45% (v/v) formamide

5× SSC

5× Denhardt

(1× Denhardt = 0.02% (w/v) BSA, PVP 360 and Ficoll 400)

— Detection of biotinylated DNA on nitrocellulose or Immobilon filters:

The filter is washed or moistened in TBST for 1 min and blocked for 1 h in TBST + 1% gelatin (Bio-Rad). It is then incubated for 20 min with a streptavidin-alkaline phosphatase conjugate (BluGene kit, BRL), 1:1000 in TBST, and then washed 2× 10 min in TBST and 2 × 5 min in AP buffer. For detection, 15 µL NBT (50 mg/L in 70% DMF, Sigma) and BCIP (25 mg/L in DMF, Sigma) in 5 mL AP buffer are added and the filter developed in it for 30 min to overnight in the dark. At longer development times the color solution (always freshly prepared) should be changed. The color reaction is terminated by washing in water and drying of the filter in air. In nitrocellulose filters, the color intensity can diminish sharply. Immobilon filters decolor slightly on rewetting in methanol.

TBST buffer: 10 mM Tris-HCl, pH 8.0

150 mM NaCl

0.05% Tween 20 (Serva)

AP buffer: 100 mM Tris-HCl, pH 9.5

100 mM NaCl

5 mM MgCl_2 in twice distilled H_2O

2.8 Preparation of affinity-purified polyclonal antibodies against T4 lysozyme

2.8.1 Immunization of rabbits with T4 lysozyme

For the injection the protein is homogenized in 0.5 mL 2× PBS and 0.5 mL Freund's adjuvant by means of ultrasound (1× PBS = 100 mM NaPO_4 , pH 7.2 and 150 mM NaCl). The first injection with 400 μg protein in 0.5 mL 2× PBS and 0.5 mL Freund's complete adjuvant is done subcutaneously into the back of the rabbit. After 3 weeks reimmunization is carried out (first booster dose) with 100 μg protein in 0.5 mL 2× PBS, 0.25 mL Freund's complete adjuvant and 0.25 mL Freund's incomplete adjuvant, also subcutaneously. The third injection (second booster dose) after another 2 weeks is administered into the thigh. It contains 200 μg T4 lysozyme in 0.5 mL 2× PBS and 0.5 mL Freund's incomplete adjuvant. 1 and 2 weeks after the second booster dose, blood is taken and tested, 7 to 15 mL of serum being obtained from 15 to 30 mL. The subsequent immunization scheme is as follows: after each second blood sample 2 weeks of rest time is allowed. After each additional booster dose (200 μg protein in 0.5 mL 2× PBS and 0.5 mL Freund's incomplete adjuvant in the thigh), one and two weeks later, blood is taken (30 mL each at a time) and a resting time of 2 weeks then interposed before the next booster dose. During each blood sampling, the blood is allowed to stand for 3 to 4 h at room temperature and the blood cake loosened from the wall and then incubated further at 4°C. The serum is transferred through a fresh tube (Falcon) and centrifuged for 15 min at 800 rpm (Heraeus Christ Centrifuge) and 4°C. In a fresh tube, it is centrifuged one more time for 15 min at 4000 rpm and 4°C. The serum is frozen at -20°C for storage.

2.8.2 Enrichment of IgG fraction from blood serum

The same volume of saturated ammonium sulfate solution in 1× PBS is added to the serum dropwise on ice during agitation (50% saturation) and further agitated for 30 min. After centrifuging, the pellet is washed once with a 50% saturated ammonium sulfate solution in 1× PBS and dissolved in the original volume of serum in 1× PBS. To eliminate ammonium sulfate, it is dialyzed extensively versus 1× PBS at 4°C.

2.8.3 Coupling of T4 lysozyme to CNBr-activated Sepharose 4B

0.5 g of the freeze-dried material (Pharmacia) is suspended in 1 mM HCl and the swollen gel (1.7 mL gel volume) washed after 15 min with 150 mL 1 mM HCl on a G3 glass frit. 8 mg T4 lysozyme are dissolved in 2.5 mL coupling buffer (0.1M NaHCO_3 , pH 8.3, 0.5M NaCl) and incubated with the prepared gel during slow rotation for 2 h at room temperature. The excess protein is washed away with the coupling buffer. The free active binding sites are saturated with

the blocking buffer (0.2M glycine, pH 8.0), for which purpose it is also slowly rolled for 2 h at room temperature. With four washing cycles of alternating pH value with 0.1M acetate buffer, pH 4.0 + 0.5M NaCl, on the one hand, and coupling buffer, on the other hand, the gel is purified and equilibrated with 0.1M Tris-HCl, pH 8.0 + 0.5M NaCl. The finished gel is stored at 4°C with 0.01% thimerosal (= Merthiolate).

2.8.4 Purification of IgG fraction on protein A-Sepharose CL-4B

The IgG molecules are selected via a protein A column from the IgG fraction (see 2.8.2) enriched from the blood serum by ammonium sulfate precipitation. 1.5 g of the freeze-dried column material (Pharmacia) are suspended in the application buffer (0.1M NaPO₄, pH 7.2) and washed with 300 mL of the same buffer after 15 min on a G3 glass frit. The swollen material is filled into a coolable column and washed with the application buffer and with elution buffer (0.1M glycine, pH 3.0) and equilibrated with the application buffer. The probe being applied is purified by centrifuging of the particles and discharged. Nonspecific material is washed out with the application buffer, then the immunoglobulin G is separated with the elution buffer. The acid eluate is immediately neutralized with 0.1 vol 1M Tris-HCl, pH 8.5. In an Amicon cell with a PM30 membrane, the eluate is desalted and rebuffered on TBS (10 mM Tris-HCl, pH 8.0, 150 mM NaCl). The protein concentration is determined according to Bradford (1976).

2.8.5 Affinity purification of anti-T4 lysozyme antibodies on Sepharose 4B-coupled T4 lysozyme

2 mg of protein A-purified immunoglobulins G (see 2.8.4) are incubated for 2 to 3 h at room temperature under slight rotation of PBS per 1 mg of the Sepharose 4B-coupled T4 lysozyme (see 2.8.3). This suspension is poured into a Pharmacia C10/20 column (4C) and washed with TBS. A small fraction of nonspecifically absorbed material is eluted with TBS + 1M NaCl. After removal of the salt by washing with TBS, the specifically bound IgG fraction is eluted with 0.1M glycine, pH 2.8. The acid eluate is immediately neutralized with 0.1 volume of 1M Tris-HCl, pH 8.5 and desalted in an Amicon cell with a PM30 membrane and rebuffered on PBS. The protein content is measured according to Bradford (1976) and set at 0.5-1 mg/mL. The affinity gel is removed from the column, washed on a G3 glass frit with 10 mM HCl (pH 2.0) and reequilibrated with TBS (addition: 0.01% thimerosal).



Figure II/1. Purification of anti-T4 lysozyme antibodies on a protein A Sepharose column, elution profile. First peak: nonspecifically bound material, second peak: specifically bound IgG antibodies.

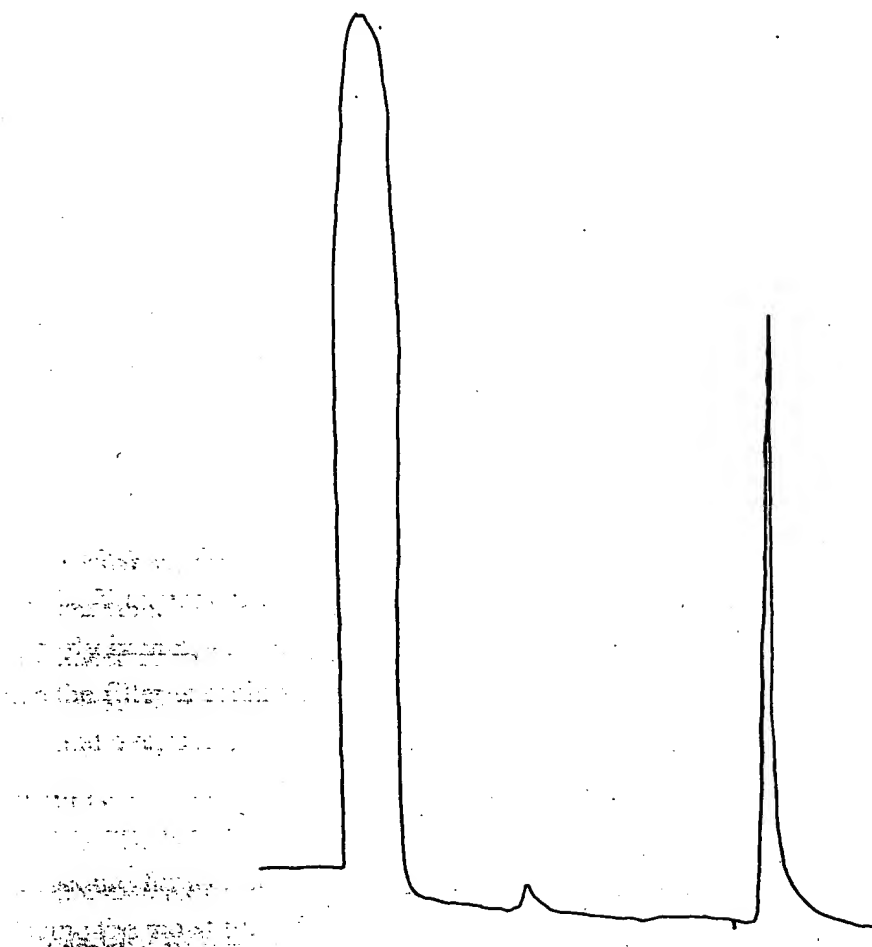


Figure II/2: Affinity purification of anti-T4 lysozyme antibodies on Sepharose-coupled T4 lysozyme, elution profile. First peak: nonspecifically bound material; second peak: specifically bound anti-T4 lysozyme antibodies.

2.9 Transformation analysis of regenerated plant material

2.9.1 Nopaline test

Detection of nopaline occurs according to Otten and Schilperoort (1978).

For callus:

About 20 mg callus material are digested in 10 μ L cold extraction buffer (0.1M Tris-HCl, pH 8.0, 0.5M sucrose, 0.1% ascorbic acid, 0.1% cysteine hydrochloride) and centrifuged. 10 μ L substrate solution (0.2M NaPO₄, pH 6.8, 60 mM arginine, 60 mM α -ketoglutarate, 16 mM NADH) are added to 10 μ L of the supernatant and incubated in the dark for 3 to 4 h. The entire charge is taken up on Whatman 3MM paper and separated electrophoretically (400 V, 90 min) using commercial nopaline (Calbiochem) as a reference. The electrophoresis buffer is formic acid/acetic acid/water (5:15:80); nopaline migrates to the cathode. After drying it is stained with

a 1:1 solution prepared from 0.02% phenanthrenequinone in ethanol and 10% NaOH in 60% ethanol and nopaline appears as a light-yellow fluorescent spot.

For plants:

The amount of nopaline formed in the plant material is sufficient in order to be detected directly. The test is therefore simplified as follows: About 10-20 mg of leaf are cut off and digested in 10 μ L cold extraction buffer and centrifuged. The supernatant is directly applied to Whatman 3MM paper. The subsequent process remains the same.

2.9.2 Neomycin phosphotransferase II test (NPT II test)

Analysis occurs according to Reiss et al. (1984) and Schreier et al. (1985).

50-100 mg plant material are homogenized in 50 μ L extraction buffer on ice. After centrifuging, the supernatant is applied to a native 10% polyacrylamide gel with collection gel. The gel is allowed to run at 70 V at 4°C. The gel is washed 2 \times 15 min in H₂O and at least 30 min in reaction buffer and then coated with an agarose gel containing NPT II substrate (1% agarose in reaction buffer, 0.1% kanamycin sulfate (Sigma) and 200 μ Ci γ -³²P-ATP). After 30 min incubation time at room temperature, a sheet of Whatman P81 phosphocellulose paper, four layers of 3MM paper and some paper towels are placed on the agarose layer and blotted for about 4 h. For reduction of the nonspecific background of transferred phosphorylated proteins, the P81 paper is treated at 65°C for 30 min with proteinase K and SDS. It is then washed 4 \times 10 min with phosphate buffer at 80°C, dried under vacuum at 80°C between filter paper and the filter exposed to x-ray film at -70°C.

Extraction buffer:	10% (w/v) glycerol
	5% (w/v) 2-mercaptoethanol
	0.1% (w/v) SDS
	62.5 mM Tris-HCl, pH 6.8
	0.025% bromophenol blue
Reaction buffer:	67 mM Tris-maleate, pH 7.1
	42 mM MgCl ₂
	400 mM NH ₄ Cl
Proteinase K solution:	2 mL 10% (w/v) SDS
	18 mL proteinase K (1 mg/mL)
Washing buffer:	10 mM NaPO ₄ , pH 7.5
Work-up of protoplasts for NPT test:	

The protoplasts are carefully centrifuged with 0.24M CaCl₂ solution, filled into an Eppendorf vessel and concentrated again briefly in a Heraeus Christ tabletop centrifuge, the

residue removed and 50 μ L cold extraction buffer added on ice, well homogenized with a small amount of sand. Subsequent analysis occurs as described.

2.10 DNA analysis of transformed plants by Southern blotting

2.10.1 Isolation of genomic plant DNA (M. Wassenegger, personal communication)

About 300-500 mg leaf are kept overnight at 4°C in 300 μ L HBT buffer and homogenized on a mortar on ice for starch degradation. The homogenate is taken up in 10 mL HBT buffer and filtered through a double layer of Miracloth (Calbiochem) in a centrifuge tube. The mortar and Miracloth are rinsed with an additional 5 mL HBT buffer and 15 mL HBT buffer is directly added to the centrifuge tube. After careful agitation, it is allowed to stand for 10 min on ice so that the chloroplasts can be exposed to the effect of Triton X-100. In a Heraeus Christ centrifuge, the nuclei are pelletized within 10 min at 3000 rpm. The supernatant is poured off, the tube allowed to drain by rotation and the remaining HBT buffer wiped dry on the edge. The pellet is resuspended in 600 μ L RBS buffer and transferred to an Eppendorf vessel. 50 μ L of a proteinase K solution (10 mg/mL in RB buffer) is added and incubated for 4 h at 37°C. This is shaken once carefully with 600 μ L phenol then again with 600 μ L phenol/chloroform/isoamyl alcohol (25:24:1). The supernatant is mixed with 0.1 vol of 3M NaAc, pH 5.0 and 2 vol warm ethanol and mixed until the polysaccharide precipitate has dissolved again and a DNA filament is visible. After 10 min of centrifuging, the pellet is washed once with warm 70% ethanol and air dried, then the DNA is dissolved overnight at 4°C in 80 μ L TE buffer.

The concentration is determined by means of A_{260} with a 5- μ L aliquot ($A_{260} = 1:50 \mu\text{g/mL}$). The genomic plant DNA is stored at 4°C. 10 μ g of DNA are digested in a total volume of 400 μ L overnight at 37°C with 100 units of restriction enzyme with the addition of RNase, extracted once with phenol/chloroform/isoamyl alcohol (25:24:1) and precipitated with NaAc and ethanol. The pellet is dissolved in 20 μ L of water.

HBT buffer:	0.3M sucrose
	5 mM MgCl_2
	50 mM Tris-HCl, pH 8.0
	1% Triton X-100
RB buffer:	30 mM Tris-HCl, pH 8.0
	10 mM EDTA
RBS buffer:	30 mM Tris-HCl, pH 8.0
	10 mM EDTA
	1% sarcosyl

2.10.2 Hybridization of filter-bound genomic plant DNA with radioactively labeled probes

The DNA digested with restriction enzyme is transferred by Southern blotting to a nitrocellulose filter as described under 2.7.4.

- Radioactive labeling of DNA fragments by oligolabeling (Feinberg and Vogelstein, 1983/1984) (see 3.4)

The probe is labeled by oligolabeling with 50 μCi $\alpha\text{-}^{32}\text{P}$ -dCTP. For this purpose, 25 ng DNA are denatured for 10 min at 95°C and cooled on ice. 1 μL of each of the dATP, dGTP and dTTP solutions is added (0.5 mM), along with the $\alpha\text{-}^{32}\text{P}$ -dCTP, 2 μL reaction mixture and 1 μL Klenow enzyme (Boehringer Random Primed DNA Labeling Kit) and allowed to react in a total volume of 20 μL for 1 h at 37°C. The reaction is stopped by 5 min of heating at 95°C.

- Hybridization with radioactively labeled probes:

The nitrocellulose filter is moistened in water and prehybridized for 2 h at 42°C in hybridization buffer (5 mL/100 cm² filter) with heat-denatured (10 min 95°C) sonicated calf thymus DNA (300 $\mu\text{g/mL}$ of buffer). It is hybridized overnight in fresh hybridization buffer (2 mL/100 cm² filter with jointly heat denatured (10 min) DNA probe and sonicated calf thymus DNA (100 $\mu\text{g/mL}$ of buffer). The filter is washed 3 \times 30 min in 3 \times SSC/0.5% SDS at 68°C, rinsed at 3 \times SSC at room temperature and air-dried. Exposure occurs overnight up to a few days.

Hybridization buffer:

50% (w/v) formamide

5 \times SSC

5 \times Denhardt

(1 \times Denhardt = 0.02% (w/v) BSA, PVP 360 and Ficoll 400)

2.11 RNA analysis of transformed plants by Northern blotting

2.11.1 Isolation of total RNA from plants (Logemann et al., 1987)

About 250 mg of leaf are homogenized in 200 μL Z6 buffer with sand and frozen in liquid nitrogen. They are then homogenized again and briefly agitated after addition of an additional 200 μL Z6 buffer. After centrifuging the supernatant is removed and the pellet reextracted with 200 μL of Z6. The combined supernatants are extracted with 1 vol phenol/chloroform/isoamyl alcohol (25:24:1) and the RNA selectively precipitated with 0.7 vol ethanol and 0.5 vol 1M acetic acid. For this purpose, it is allowed to stand at -20°C overnight. The RNA pellet is washed once with cold 3M NaAc, pH 5.0 and once with cold 70% ethanol. The liquid is carefully removed and the RNA taken up in 100 μL H₂O. If the RNA is not fully dissolved at room temperature, it is briefly heated to 56°C. With a 5- μL aliquot, the concentration is determined via A₂₆₀ (A₂₆₀ = 1:40 $\mu\text{g/mL}$). The RNA solution is stored at -20°C.

Z6 buffer:

8M guanidine hydrochloride, pH 7.0

20 mM MES
 20 mM EDTA
 50 mM 2-mercaptoethanol

2.11.2 RNA formaldehyde agarose gels

50 µg total RNA are precipitated from the optically measured solution (see 2.11.1) with 0.1 vol 3M NaAc, pH 5.0 and 2 vol ethanol, centrifuged and the liquid well removed. The pellet is briefly dissolved in 20 µL H₂O at 56°C, denatured with 20 µL formamide, 6 µL 37% formaldehyde and 4 µL 10× MEN buffer 15 min at 56°C. 5 µL blue marker are then added to the coating for this purpose. For a formaldehyde-agarose gel, 6 g agarose are dissolved in 300 mL H₂O and cooled to 60°C. After addition of 65 mL 37% formaldehyde and 40 mL 10× MEN buffer a bridge gel is cast and cooled at 4°C. After heating to room temperature, the probes are loaded and the buffer reservoir filled with 1× MEN buffer. The gel runs 30 mA (120 V) for about 4 to 5 h. When the probes are run, the coating wells are filled with 1× MEN buffer and the gel surface covered air bubble-free with preserving film. After each hour the contents of the two buffer reservoirs are mixed in order to prevent formation of a pH gradient. The gel can be stained in 500 mL 50 mM NaOH with 200 µL ethidium bromide (10 mg/mL) for 30 min at room temperature. It is decolorized by repeated washing in 10 mM phosphate buffer, pH 7.0 after 4 to 5 h and the RNA is visible under UV light.

10× MEN buffer:

10 mM MOPS, pH 7.0

50 mM NaAc

10 mM EDTA

2.11.3 Hybridization of filter-bound RNA with radioactively labeled probes

RNA-formaldehyde gels are blotted without any additional pretreatment. Transfer to the nitrocellulose or Immobilon membrane occurs as described for the Southern blots (see 2.7.4). After completed transfer, the filter is air-dried without washing (5 min) and baked in a vacuum furnace for 2 h or fixed for 4 min on a transilluminator. Hybridization occurs under the same conditions as for the Southern blot (see 2.10.2).

2.11.4 Isolation of total RNA from protoplasts

For isolation of RNA from protoplasts, these are resuspended after mild centrifuging in 200 µL Z6 buffer and frozen in liquid nitrogen. After thawing on ice, RNA isolation begins as described under 2.11.1 with homogenization.

2.12 Protein analysis of transformed plants

2.12.1 Tissue printing (Cassab and Varner, 1987)

The rapid method serves to perform immunodetection with plant tissue directly without producing an extract. For this purpose, nitrocellulose is placed for 30 min in 0.2M CaCl_2 and then dried on paper. Immobilon PVDF (Millipore) is wetted in methanol and substituted in TBST. The tissue is cut and placed with the cut surface on the prepared membrane and additionally pressed against the filter for 30 sec. It is then dried with hot air (in nitrocellulose) and detection begins as described in Western blotting in 2.12.10. In this method, which uses one spot from the total protein as the probe, it can be particularly important to initially incubate the employed antibody with wild type total protein extract in order to avoid cross reaction.

2.12.2 Isolation of total protein from plants

For probe preparation for direct coating on a denaturing SDS-PAA gel, the plant tissue is cut and mixed with 10 μL cold 2 \times SDS probe buffer/10-30 mg and homogenized while cold. The cell fragments are centrifuged, the volume made up with H_2O to 1 \times SDS probe buffer concentration (total 20 μL) and the supernatant heated for denaturation of the proteins before coating for 5 min at 95°C. It is then centrifuged one more time. If protein determination is to be conducted (Bradford, 1976) a PBS extract (0.1M NaPO_4 , pH 7.2, 150 mM NaCl) must first be prepared in the same manner, since the SDS and DTT fractions interfere with these measurements in the SDS probe buffer. It is then made up with 2 \times SDS probe buffer. To inhibit proteases, a combination of protease inhibitors must be added in this case (200 μM PMSF, 1 μM leupeptin, 1 μM pepstatin, 100 μM EDTA).

1 \times SDS probe buffer:

75 mM Tris-HCl, pH 6.8

1% SDS

10% glycerol

0.002% bromophenol blue

100 mM DTT

2.12.3 Isolation of total protein from protoplasts

After isolation of protoplasts from the culture medium, these are frozen in 10 μL 2 \times SDS probe buffer in liquid nitrogen. After thawing on ice, the protoplasts are homogenized with a small amount of sand, centrifuged and the supernatant adjusted to 1 \times SDS probe buffer. Before application to the polyacrylamide gel, it is denatured for 5 min at 95°C.

2.12.4 Isolation of protein from protoplast culture medium

The protoplasts are centrifuged with the smallest possible volume of 0.24M CaCl_2 solution and the supernatant collected. These are introduced in 2-mL aliquots to the Centricon 10 units (Amicon) and centrifuged at 4500 G and 10°C for about 2 h until a residual volume of about 50 μL is reached. The combined concentrates of the aliquots enriched by ultrafiltration are rebuffered in the same manner once without 1.5-2 mL PBS buffer. The concentrate is collected in 2 min at 900 G and concentrated to about 10 μL in the Speed-vac. After addition of 10 μL 2 \times SDS probe buffer and heat denaturation (5 min 95°C), the probe is applied to an SDS-PAA gel.

2.12.5 $(\text{NH}_4)_2\text{SO}_4$ precipitation of T4 lysozyme plant extract

2 g of leaves of a transformed plant are homogenized in a mortar in 2 mL PBS (0.1M NaPO_4 , pH 7.2, 150 mM NaCl) with protease inhibitor (PI) (200 μM PMSF, 1 μM leupeptin, 1 μM pepstatin, 100 μM EDTA) and this washed with 8 mL PBS with PIs. After centrifuging, the pellet is washed with 3 mL PBS with PIs and an $(\text{NH}_4)_2\text{SO}_4$ precipitation is carried out. The 20-40% saturated fractions are collected. For this purpose, 250 μL of saturated ammonium sulfate solution in PBS/1 mL of extract are added on ice during agitation (20% saturation) dropwise and further agitated for 30 min for equilibration. After centrifuging of the pellet, an additional 416 μL of saturated solution/mL of extract are added dropwise to the supernatant (40% saturation). The precipitated proteins are collected by centrifuging and the pellet washed with 40% saturated $(\text{NH}_4)_2\text{SO}_4$ solution in PBS. This protein fraction is dissolved in 2 mL PBS and extensively dialyzed versus PBS at 4°C.

2.12.6 Coupling of polyclonal anti-T4 lysozyme antibodies to CNBr-activated Sepharose 4B

As described under 2.8.3, 200 μg of the affinity-purified (see 2.8.5) polyclonal anti-T4 lysozyme antibody is coupled to 0.4 g of the activated Sepharose. The finished material is suspended in a final volume of 3.5 mL, 50 μL of which is sufficient for up to 500 mg of leaf material.

2.12.7 Affinity purification of exogenous proteins from plant extract

The plant material is homogenized in a mortar during cooling with liquid nitrogen. The homogenate is filled into a Corex tube and the extraction buffer and PVPP added. 50 mg PVPP and 1 mL of 1 \times PBS with protease inhibitors (PIs) (200 μM PMSF, 1 μM leupeptin, 1 μM pepstatin, 100 μM EDTA) are added per gram of tissue. As protection against oxidizing substances in the plant extract, 5 mM ascorbic acid is additionally introduced. In order to solubilize all the proteins present, i.e., even those enclosed by membranes, a detergent must be added. 1% Triton X-100 or preferably 10 mM CHAPS is suitable for this. After 15 min of extraction, the cell fragments are centrifuged at 10,000 rpm at 4°C. When a detergent is used, the

extraction solution for affinity chromatography must be diluted to about 0.1% or 2 mM detergent. The supernatant is rolled on a Denley roller at 4°C to precipitate supersaturated oligosaccharides from 30 to 60 min and then centrifuged one more time at 4000 rpm and 4°C. The total amount of protein is equalized by a Bradford protein determination. The total protein solution is mixed with the affinity gel and incubated overnight at 4°C with slight rolling. The affinity gel is centrifuged at 4000 rpm and 4°C, the supernatant is discarded and the gel transferred with a small amount of buffer to an Eppendorf vessel. In this vessel, it is washed cold once with 1 mL 1× PBS with PIs, once with 1 mL 1× PBS/1M NaCl + PIs and 3× with 1 mL 1× PBS + PIs and the bound specific protein eluted with 0.1M glycine, pH 2.8 or for NP Sepharose with 10^{-4} M NIP-cap (3× 200 μ L). The acid eluate is immediately neutralized with 0.1 vol 1M Tris-HCl, pH 8.0. The eluate is introduced to the Centricon 10 unit (Amicon) and concentrated by ultrafiltration and then desalted with H₂O and concentrated to about 50-75 μ L. The volume of the probe is further reduced to about 10 μ L in the Speed-vac. For this purpose, the same amount of 2× SDS probe buffer is added and the probe applied to a Western gel.

2.12.8 Western blotting

The prepared denatured probes are separated on a denaturing 10-15% discontinuous SDS polyacrylamide gel (system according to Laemmli, 1970). A minigel apparatus (8 × 8 cm) is used, and the gels are run under standardized conditions:

For lysozyme analyses: 15% gel with 20 mA/max 200 V, 90 min

For antibody analyses:

Light chain: 15% gel with 20 mA/max 200 V, 120 min

Heavy chain: 10% gel with 20 mA/max 200 V, 70 min

The principles of Western blotting are described in Towbin (1979) and Burnette (1981). A review is offered by Gershoni and Palade (1983).

A new development of Western blotting, so-called semidry blotting, is used here (Kyhse-Andersen, 1984). An in-house apparatus is used, consisting of two carbon electrode plates mounted on a plastic support. Introduction occurs precooled in a cool room at 4°C. After the run, the polyacrylamide gel is equilibrated for 30 min at room temperature in the blotting buffer. The blot is constructed as follows: the lower electrode is the cathode, three layers of Whatman 3MM paper, which are immersed in the blotting buffer are placed on it cut to the correct gel size and then the gel is positioned free of air bubbles and then a nitrocellulose filter (Schleicher & Schuell BA 85, 0.45 μ m) in gel size or a hydrophilic Immobilon PVDF membrane (0.45 μ m, Millipore) and another three immersed 3MM papers. The seal is formed by the anode plate. The entire structure must be free of air bubbles in order to achieve uniform transfer. In order to achieve uniform contact of the electrode plates, a weight should be placed on the center

of the blot structure. Electrode transfer occurs in 25 min at 200 mA. Slight heating is found in this period. The efficiency of transfer, however, can be checked by Coomassie or silver staining of the gel. A nitrocellulose filter or the Immobilon membrane is stained with Ponceau S (Serva) and the size standard marked. It is largely decolored with water. The nonspecific binding sites are blocked for immunodetection for 1 h with TBST buffer + 1% gelatin (Bio-Rad) at room temperature. The filter is washed briefly in TBST and then incubated with the specific first antibody in appropriate dilution in TBST for 2 h to overnight. After 10 min of washing three times in TBST, it is incubated for 45 min with biotinylated second antibody at the appropriate dilution in TBST. It is again washed for 3× 10 min with TBST. It is then incubated for 20 min with a streptavidin-alkaline phosphatase conjugate (BRL, from the BluGene Kit), washed for 2× 10 min in TBST and 2× 5 min in AP buffer. For detections, 15 µL NBT (50 mg/mL in 70% DMF, Sigma) and BCIP (25 mg/mL in DMF, Sigma) in 5 mL AP buffer are added and the filter developed in it for 5 min to 2 h in the dark. In special cases, it is also worthwhile to continue to develop overnight which, however, results in a stronger background. The color solution should always be used fresh. The reaction is terminated by brief washing in water and drying in air. Storage is done in the dark. The original sensitivity can be restored in the nitrocellulose filter by moistening. The Immobilon membrane can be partially decolored by wetting in methanol.

In order to reduce nonspecific interactions of the employed first and second antibodies, washing can first be carried out once with PBS/1M NaCl and then three times with PBS. The incubations can also be carried out with 0.1% Triton X-100, 0.1% Tween-20 and 0.3% gelatin.

Blotting buffer:

20 mM Tris-HCl

150 mM glycine

10% (w/v) methanol

TBST buffer:

10 mM Tris-HCl, pH 8.0

150 mM NaCl

0.05% Tween 20 (Serva)

AP buffer:

100 mM Tris-HCl, pH 9.5

100 mM NaCl

5 mM MgCl₂

2.12.9 Employed antibodies for immunodetection

First antibody:

Anti-T4 lysozyme	1:1000 (rabbit/polyclonal)
Anti-light chain	1:1000 (goat/polyclonal)
Anti-IgM (heavy chain)	1:1000 (goat/polyclonal)
AF6 (anti-IgM)	1:1000 (mouse/monoclonal)
Ac 38 (anti-B1-8)	1:1000 (mouse/monoclonal)
Ac 164 (anti-B1-8)	1:1000 (mouse/monoclonal)
LS 136 (anti- α)	1:1000 (mouse/monoclonal)

Reference for Ac 38, Ac 146, Ls 136: Reth (1981).

Second antibody:

Anti-rabbit IgG, biotinylated	1:2000 (goat) (Bio-Makor)
Anti-goat IgG, biotinylated	1:2000 (rabbit) (Sigma)
Anti-mouse IgG, biotinylated	1:2000 (goat) (Sigma)

2.12.10 Silver staining of protein polyacrylamide gels

After the run, the gel is fixed for 20 to 30 min in 10% acetic acid and 25% propanol. After treatment with 10% glutaraldehyde (20-30 min) it is washed during changing with water. This is followed by incubation with 0.03 mM DTT for 20 to 30 min and staining with 0.1% silver nitrate also from 20 to 30 min. The bands are made visible in about 5 min with the development solution (30 g NaCO_3 and 500 μL 37% formaldehyde per liter). The reaction is stopped with 10% acetic acid/25% propanol. If the gel is to be dried, 1% glycerol is added in the last washing step.

2.12.11 T4 lysozyme activity test with *Micrococcus lysodeicticus*

(modified according to Selsted and Martinez, 1980, see also McKenzie and White, 1986)

A fresh suspension in phosphate buffer (pH 7.2-7.2) is produced from lyophilized *Micrococcus lysodeicticus* so that an A_{450} value of 0.6 is obtained at 1:10 dilution. 0.1 mL of the cell suspension and up to 0.1 mL of the probe being analyzed are combined for the reaction, adjusted to 0.1% NaN_3 , 200 μM PMSF, 1 μM leupeptin, 1 μM pepstatin, 100 μM EDTA and 1 mg/mL BSA and made up to 1 mL with phosphate buffer (pH 7.2-7.4). This mixture is incubated for 18 h at 37°C during slow rolling. A control is run without the probe liquid. A_{450} is then measured and the "averaged" reduction of the absorption is plotted versus the logarithm of lysozyme concentrations and compared with the control value. A linear region with significant reduction of absorption begins in the measured standard affinity curve in concentrations from about 100 ng/mL upward.

2.13 Immunogold labeling

2.13.1 High pressure freezing preparation

Native tissue without chemical pretreatment is used for the preparation of plant material. The biologically appropriately prepared material is cut to fit and then shock-frozen ultrafast in liquid nitrogen on a Balzers HPM 010 (kindly furnished by H. Moor, Zurich). Copper plates or preferably gold plates cured in hydrogen peroxide can be used as preparation holders.

Storage of the samples occurs in liquid nitrogen.

2.13.2 Freezing substitution

Freezing substitution in 0.5% uranyl acetate in anhydrous acetone in the following steps: -90°C (8 h), -60°C (overnight), -35°C (5 h). All the following steps are conducted at -35°C: two times 1 h of rinsing with acetone and then first infiltration with Lowicryl HM 20 1:1 mixed with acetone for 4 h, then Lowicryl HM 20/acetone 2:1 overnight; then pure Lowicryl HM 20 over the day time with two-fold changing and again overnight, on the next day the infiltration is concluded one more time for 5 h with fresh Lowicryl HM 20, the samples are filled into polymerization capsules (Beem capsules) and the capsules filled with fresh Lowicryl HM 20. Polymerization of the synthetic resin occurs overnight in indirect UV light (360 nm) and then for another 8 h with direct UV exposure in each case at -35°C. The cured samples are brought to room temperature.

2.13.3 Ultramicrotomy

Silver sections (about 60 nm) are prepared on an ultracut microtome from the Reichert Co. by means of a diamond blade. These are trapped on pioloform-filmed and carbon-sputtered nickel grids (100 mesh).

2.13.4 Immunocytochemistry

The sections are initially moistened in PBS (0.1M, pH 7.2) for 5 min. For saturation of nonspecific binding sites both in plant tissue and in the film coating of the grid, they are incubated in alternation twice with 0.02M lysine in PBS and with 0.1% gelatin in PBS first for 10 min, then 5 min. This is followed by further blocking for 10 min in PBS with 1% BSA, 0.1% Triton X-100 and 0.1% Tween 20 (incubation buffer). The reaction with the first antibody (50-100 µg/mL) in incubation buffer occurs over 2 to 3 h at room temperature. It is then washed for 5 min in incubation buffer and 25 min in PBS with 0.5% Tween 20. The first antibodies are described under 2.12.9. Incubation with a second antibody (biotinylated) occurs in the same manner. As reagent absorbed on colloidal gold for direct reaction with the first antibody, protein

A (Janssen/Auroprobe EM protein A/G 10) for polyclonal antibodies, anti-mouse IgG from goats (Janssen/Auroprobe EM goat anti-mouse IgG/G 10) for monoclonal antibodies or streptavidin (Janssen/Auroprobe EM streptavidin/G 10) for biotinylated antibodies are used. The immunogold reagent is diluted 1:10 in the incubation buffer and applied for 1 h at room temperature. It is then rinsed for 15 min in PBS with 0.5% Tween 20 and 15 min in twice distilled water and the preparation dried. All employed antibodies are preincubated before addition of the preparation for 10 min with a W38 wild type PBS extract with addition of protease inhibitors (see 2.12.8) in order to saturate the nonspecific binding sites.

Analysis occurs on an EM 10 (Zeiss) at 60 keV on uncontrasted probes and post-contrasted preparations. For contrasting, incubation is carried out for 15 min in 8% uranyl acetate and for 15 min in lead nitrate according to Reynolds (1963).

III. Results

3.1 Construction of a T_R promoter-NOS promoter vector with chimeric genes for α -amylase signal peptide and light chain or heavy chain of the antibody

The final selection-expression vector for production and secretion of a monoclonal antibody in plants was assembled in several steps from different components. In the first step, the signal peptide gene of α -amylase from barley was subcloned (Rogers and Milliman, 1983) from the clone E (obtained from J. C. Rogers) (see Figure III/3).

Clone E contains an α -amylase (cDNA) of the Pst I cleavage site of pBR 322. The coding region begins in an Nco I cleavage site that contains an ATG-codon so that during the entire cloning work, a check for obtaining a correct translation start is possible in a simple fashion. The processing site and thus the beginning of the mature protein lie after amino acid 23 (serine) of the signal peptide. There is no suitable cleavage site here with which the sequence could be cut after the serine codon. An adjacent Bgl I cleavage site is used.

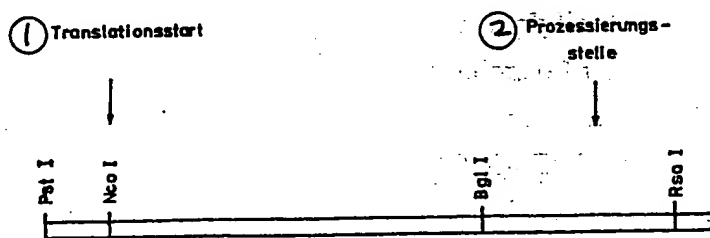
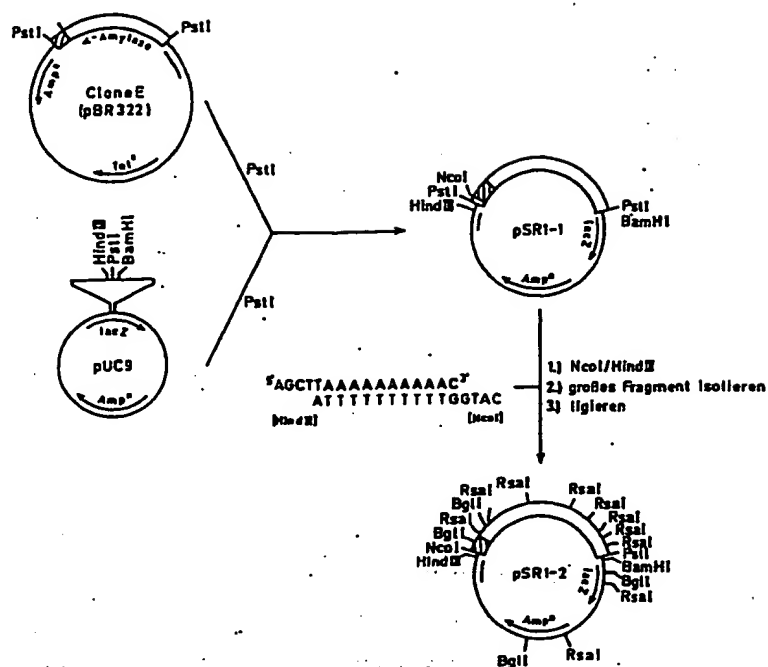


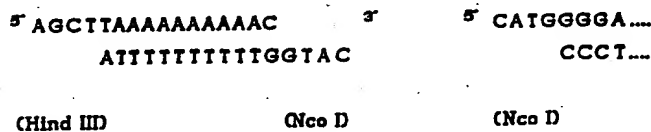
Figure III/1: Structure of α -amylase signal peptide-coding gene in clone E (Rogers and Milliman, 1983).

Key: 1 Translation start

2 Processing site

Figure III/2: Cloning of α -amylase signal peptide gene.

Key: 1 Nco I/Hind III
 2 Isolate large fragment
 3 Ligase

Figure III/3: Structure of synthetic Hind III linker oligonucleotide and the 5' region of the α -amylase signal peptide gene.

Initially the Pst I-cDNA clone was cut out from pBR 322 and cloned into the Pst I cleavage site in the polylinker of the pUC 9 (pSR 1-1) so that a Hind III cleavage site lies

directly adjacent 5' to the Pst I cleavage site (see Figure III/2). Since all constructs should be cloned originally in a Hind III cleavage site of a cassette, this was ligated to the Nco I cleavage site as a 5' end by a synthetic oligonucleotide (synthesis and purification, see 2.7.3).

The Hind III and Nco I cleavage sites were bound by an oligo-dA/dT strand in order to increase the binding capacity of the proteins to the DNA double strand. It is known that AT-rich regions readily melt and in so doing permit binding of proteins (Salomon et al., 1986; Joshi, 1987) and this facilitates the function of the 40S ribosomal subunit for translation initiation (Heidecker and Messing, 1986; Kozak, 1986). The plasmid pSR 1-1 was digested with Hind III and Pst I, the large fragment isolated and the synthetic oligonucleotide ligated in (pSR 1-2). In this form, the signal peptide gene is readily isolatable and for this purpose 5' the Hind III and 3' the Bgl I cleavage sites are used (see Figure III/1). The missing nucleotides must be replaced by synthetic oligonucleotides.

For construction of a chimeric signal peptide-light chain-gene (see Figure III/4), the gene for the light chain was excised from the vector plasmid pLGV 2385 Lc (Steiger, 1987) by combined Xho II/Hha I digestion and thus ligated into the Bam HI cleavage site of pBR 322 so that the 3' end came to lie in the Tet^R gene (pSR 1-10). The plasmid was checked by Southern hybridization with the gene for the light chain, in addition to restriction mapping. The 3' end of the gene for the light chain and the Tet^R were excised with Pvu II and pSR 1-12 formed by self-ligation. The 5' part of the gene for the light chain is now present in this plasmid, in front of which the signal peptide gene can be ligated. By restriction digestion with Rsa I, a 1.87 kb fragment is isolated from plasmid pSR 1-2, which contains the signal peptide gene (see Figure III/5). Of the four Bgl I cleavage sites originally contained in pSR 1-2, only two therefore still remain. This fragment is shortened to 1.45 kb by digestion with Bgl I. A synthetic oligonucleotide was ligated to it, which contains the 3' part (the missing nucleotide) of the signal peptide and the first nucleotide of the 5' end of the gene for the light chain (see Figure III/6). After digestion with Hind III, in addition a small fragment (86 bp) which represents the signal peptide gene, an additional large fragment (1.35 kb) is obtained. This fragment mixture is ligated in Bam HI/Hind III-opened pSR 1-12. The clone that contains the correct (small) fragment with the signal peptide gene can be easily identified with reference to the size difference and by Nco I digestion of DNA minipreparations (pSR 1-3). From pSR 1-3 the 5' part of the chimeric gene is isolated with part of the plasmid by digestion with Ava II and Pvu I and ligated with the 3' part isolated from pLC 1-4 (Stieger, 1987) by digestion with Ava II and Hind III, in which, due to the present restriction cleavage sites, only the correct arrangement can form. By digestion with Hind III, the complete chimeric gene is liberated and ligated in the Hind III cleavage site of pUC 9 (pSR 1-14). The terminal cleavage sites are converted by ligation of a Sal I linker to Sal I cleavage sites (pSR 1-16). In the last step, the chimeric signal peptide-light chain-gene isolated

with Sal I is ligated in the correct orientation in the Sal I cleavage site of the vector plasmid pAP 2034 (Velten and Schell, 1986) (pSR 1-4). This selection-expression vector contains the agrobacteria- T_R double promoter (Velten et al., 1984). The cassette on the 1' side contains singular Sal I and Bam HI cloning sites and the polyadenylation signal of gene 7 of the T-DNA (g7pA). The NPT II gene of Tn5 is incorporated as a selection (kanamycin and G 418 resistance) and marker gene for plants in the 2' position of the promoter, provided with the polyadenylation signal of octopine synthase gene (ocs pA). Polyadenylation is a process in eukaryote cells that stabilizes the mRNA in cytoplasm and favors translation of mRNA on ribosomes (Wickens and Stephenson, 1984; Birnstiel et al., 1985). In addition, the plasmid pAP 2034 contains the antibiotic resistances against carbenicillin and streptomycin/spectinomycin, which are necessary for selection of transconjugants between *E. coli* and *A. tumefaciens*. Different restriction digestions for control of plasmid pSR 1-4 are shown in Figure III/7.

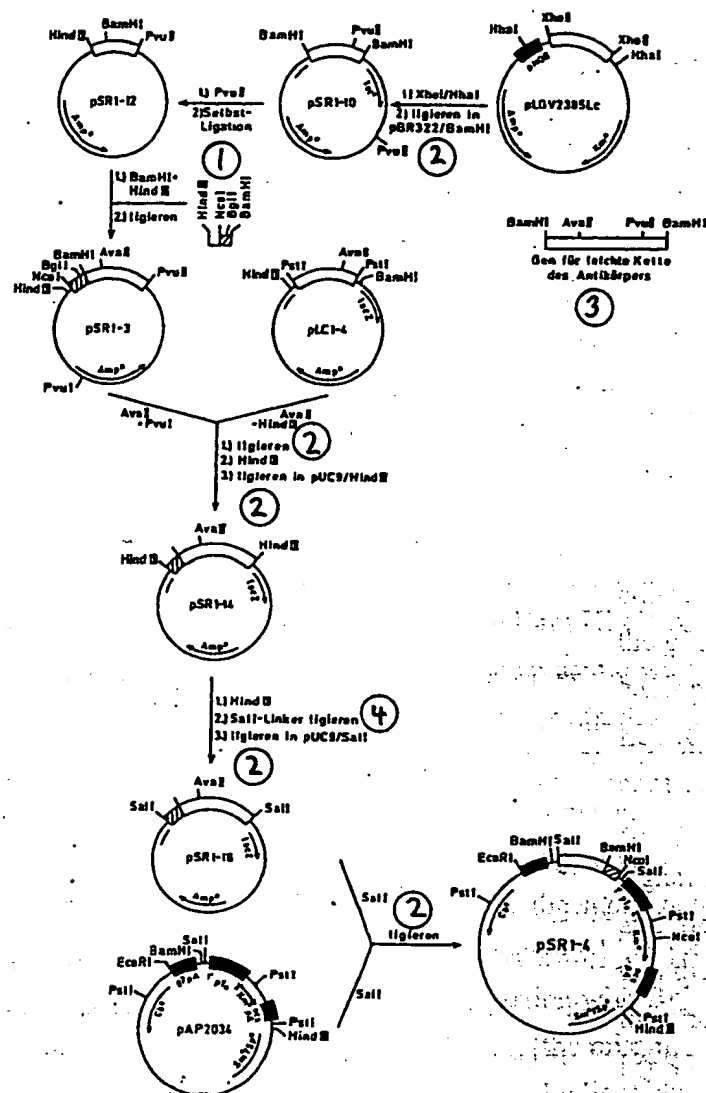


Figure III/4: Construction of a chimeric α -amylase signal peptide-light chain gene under control of the T_R promoter.

- Key:
- 1 Self-ligation
 - 2 Ligate
 - 3 Gene for light chain of the antibody
 - 4 Ligate Sal I linker

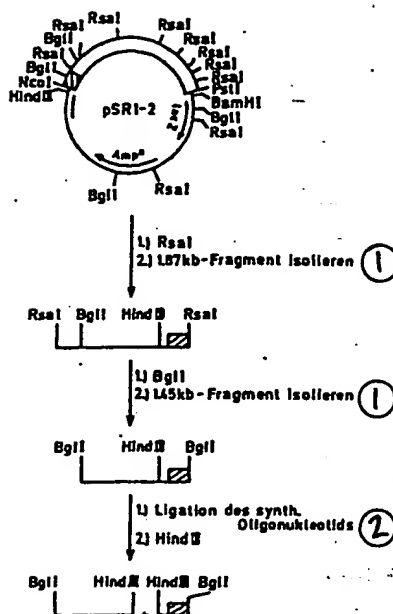


Figure III/5: Isolation of the signal peptide gene.

Key: 1 Isolate fragment
2 Ligation of synthetic oligonucleotide

5' TGGCGTCC|G 3' start signal peptide
CCAACCGCAGG|CCTAG
Bgl II Bam HI

Figure III/6: Synthetic oligonucleotide for joining of the signal peptide and light chain gene.

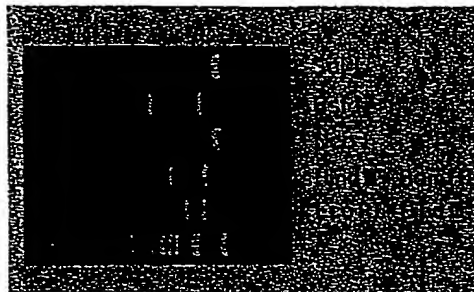


Figure III/7: Checking of plasmid pSR 1-4 by restriction mapping.

The plasmid DNA was digested with the mentioned enzymes and separated on a 0.7% TBE-agarose gel. λ -DNA digested with Pst I serves as length standard.

Fragment sizes: Sal I: 8.6/0.9 kb, Pst I: 5.4/1.63/1.55/0.97 kb, Nco I: 8.4/1.1 kb, Hind III + Bam HI: 6.37/2.38/0.8 kb, Eco RI + Hind III: 6.15/3.4 kb.

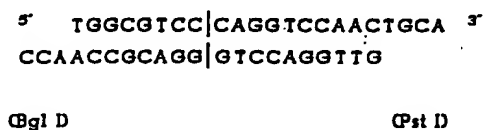


Figure III/8: Synthetic oligonucleotide for joining of the signal peptide in heavy chain gene.

Parallel with construction of the chimeric light chain gene, a similar signal peptide-heavy chain gene fusion is constructed (Figure III/9). For this purpose, the signal peptide gene is first isolated as described (see Figure III/5) from the plasmid pSR 1-2 and ligated similarly with a suitable synthetic oligonucleotide that contains the missing nucleotide of the signal peptide gene and the first nucleotide of the 5' part of the gene of the heavy chain up to the Pst I cleavage site (see Figure III/8). The signal peptide gene is then flanked 5' by a Hind III and 3' by a Pst I cleavage site. This fragment is cloned into the plasmid pHG 1-3 (Stieger, 1987) opened with Pst I and Hind III, which contains the 5' part of the gene for the heavy chain and in so doing fusion occurs between the two genes (pSR 3-1). The 3' end of the heavy chain gene is excised from

plasmid pHC 3-19 (Stieger, 1987) with Bam HI and Hind III, ligated with the 5' part of the fusion gene (Hind III/Bam HI fragment from pSR 3-1) and then cloned into the pUC 9 plasmid opened with Hind III (pSR 3-10). This plasmid contains the complete chimeric gene flanked by Hind III cleavage sites. After its isolation, a similar pUC 9 derivative with the Sal I-flanked chimeric gene is produced by fill-in and ligation of a Sal I linker (pSR 3-12). For closure, an expression vector must be constructed that contains the two chimeric genes under the control of regulation units active in plants in order to be able to simultaneously transfer both genes (see Figure III/11). As the only selectable marker at the time of construction, the neomycin phosphotransferase II gene with kanamycin or G 418 resistance of the transformed plants was useful. It was therefore only possible to select for the transformation with a plasmid or a T-DNA. For expression of both chains of the antibody, these genes therefore had to be incorporated in the same plasmid under two different promoters in order to obtain a stable arrangement.

For this purpose, the plasmid pSR 1-4 is chosen as the starting material (with the chimeric gene for the light chain) and cloned into the singular Eco RI cleavage site in addition to the g7pA region of the chimeric gene for the heavy chain. Initially, a fragment that contains the NOS promoter (Depicker et al., 1982, Bevan et al., 1983, Sanders et al., 1987) is isolated with Eco RI and Hind III from the plasmid pLGV 2385 Bgl → Hind (M. Stieger, personal communication), in which the Bgl II cloning site was converted into a singular Hind III cloning site (with elimination of an originally present Hind III cleavage site). The 5' part of the chimeric signal peptide-heavy chain gene is excised from the plasmid pSR 3-1 with Hind III and Bam HI and the two fragments subcloned in pUC 9 opened with Eco RI and Bam HI (pSR 4-1). From this the promoter gene fragment can be excised again with Eco RI and Bam HI and ligated together with the Bam HI-Sal I-3' part of the coding region in pUC 9 opened with Eco RI and Sal I (pSR 4-2). The 3' fragment is excised beforehand from the plasmid pSR 3-12 with Bam HI and Sal I. With the same strategy, for completion of gene construction and the expression vector, the g7pA fragment (Sal I-Eco RI) from pAP 2034 is ligated to the 3' end of the chimeric heavy chain gene and the entire expression unit cloned into the singular Eco RI cleavage site of plasmid pSR 1-4 in an orientation so that the NOS promoter comes to lie adjacent to an already present g7pA signal and the two homologous regions that are roughly 300 bp long are separated by a roughly 2-kb nonhomologous sequence.

The finished selection-expression vector for the genes for the light and heavy chains of the monoclonal B1-8-IgM antibody (pSR 4-3) contains the NPT II gene (with ocs pA) under the control of the T_R double promoter for selection and the chimeric gene for the signal peptide-light chain fusion (with g7pA), under the control of the NOS promoter and the g7pA region of the chimeric gene for the signal peptide-heavy chain fusion and the carbenicillin as well as

streptomycin/streptomycin resistance gene. The entire plasmid has a size of 11.8 kb and checking occurs by numerous restriction digestions (Figure III/12).

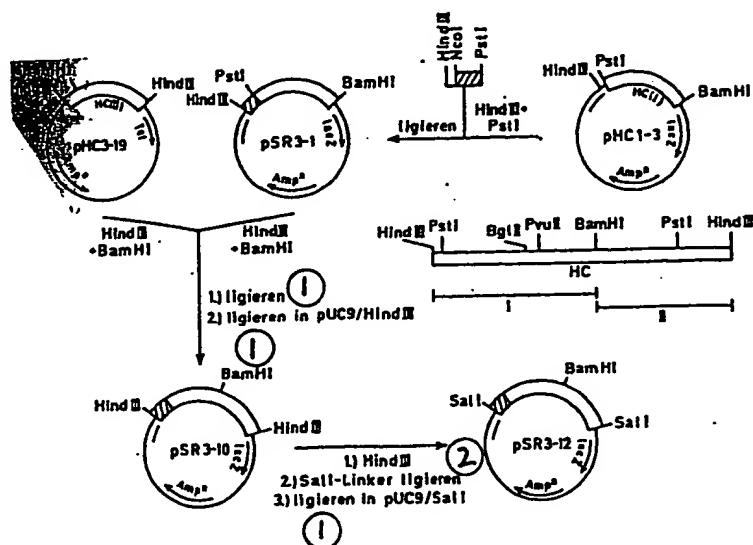


Figure III/9. Construction of a chimeric α -amylase signal peptide heavy chain gene.

Key: 1 Ligate
2 Ligate Sal I linker

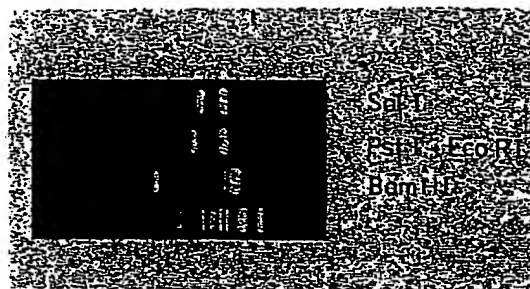


Figure III/10. Checking of plasmid pSR 3-12 by restriction mapping.

The plasmid DNA was digested with the mentioned enzymes and separated on a 0.7% TBE-agarose gel. λ -DNA digested with Pst I serves as length standard.

Fragment sizes: Sal I: 2.7/1.7 kb, Pst I + Eco RI: 2.7/1.5/0.1/0.1 kb,

Bam HI: 3.5/0.8 kb.

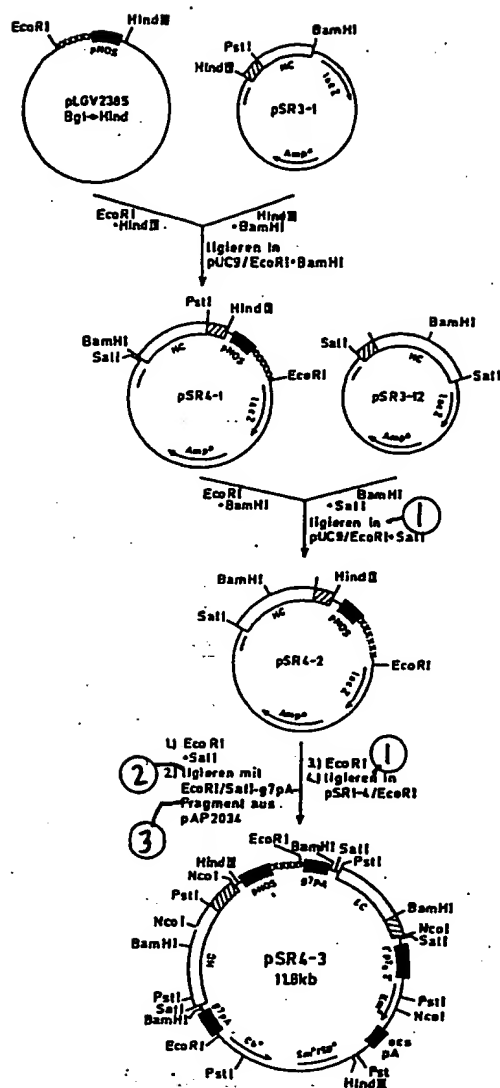


Figure III/11. Construction of a selection-expression vector for the chimeric α -amylase signal peptide-light chain or heavy chain gene under control of T_R or NOS promoter.

Key: 1 Ligate
 2 Ligate with
 3 Fragment from

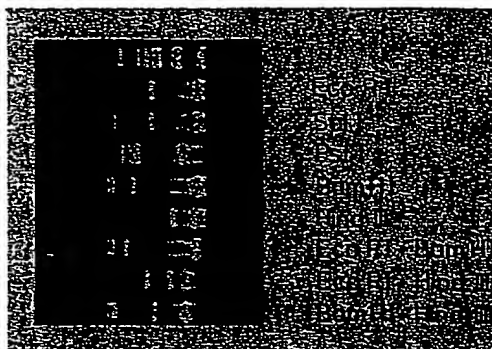


Figure III/12. Checking of plasmid pSR 4-3 by restriction mapping.

The plasmid DNA was digested with the stated enzymes and separated on a 0.7% TBE-agarose gel. λ -DNA digested with Pst I serves as length standard.

Fragment sizes: Eco RI: 9.5/2.3 kb, Sal I: 8.6/2.3/0.9 kb, Pst I: 5.4/1.63/1.45/1.45/1.16/0.64 kb, Bam HI: 8.7/1.53/0.83/0.78 kb, Hind III: 8.1/3.75 kb, Eco RI + Bam HI: 8.5/1.29/0.83/0.78/0.22/0.22 kb, Eco RI + Hind III: 6.15/3.4/1.96/0.33 kb, Bam HI + Hind III: 6.37/2.37/0.96/0.83/0.78/0.55 kb.

3.2 Construction of the T_R promoter vector with the chimeric gene for α -amylase signal peptide and T4 lysozyme

The chimeric signal peptide-T4 lysozyme gene is incorporated in the free 1' position of the T_R double promoter in the vector pAP 2034 (Velten and Schell, 1986). As described in Figure III/5, the signal peptide gene, during ligation of the synthetic oligonucleotide shown in Figure III/13 that includes the last nucleotide of the signal peptide gene and a larger fraction of the beginning of the T4 lysozyme gene, is isolated. The oligonucleotide is provided on the 3' end with a Bam HI cleavage site as a cloning aid and in addition it contains an Sna BI cleavage site for ligation to the Sna BI cleavage site contained in the gene. This oligonucleotide is ligated into the pUC 9 plasmid cleaved with Hind III and Bam HI (pSR 2-1) and prepared by Sna BI and Eco RI treatment of the recombinant plasmid for take-up of the T4 lysozyme gene (see Figure III/14).

The originally prescribed strategy of direct cloning of the entire T4 lysozyme gene that has a 3.2-kb Hind III fragment from the DNA of phage λ 1358-b 538 imm:T4 EIII e^+ sus97

(Owen et al., 1983) did not lead to success either with pBR 322 or with pUC 9. Instead, even incorporation of the λ arm in the plasmid was observed, but not that of the Hind III-T4 lysozyme fragment; this can be explained by possible lytic activity of the gene products of this DNA section so that no bacterial clones could be cultured. The isolated λ DNA is digested with Hind III and the 3.2-kb fragment isolated (see Figure III/14). After an additional digestion with Ava II, the coding fragment reduced to about 750 bp is isolated. Starting from here, the T4 lysozyme gene is cloned in two parts so that at no point is the complete gene without signal peptide gene fusion under control of a possibly active promoter. In the prepared plasmid pSR 2-1, the roughly 200-bp large Sna BI/Eco RI 5' fragment is ligated (pSR 2-2) and after removal of the Eco RI/Nde I fragment of the lac Z gene from the plasmid, the roughly 350-bp-long Eco RI/Nde I 3' fragment (pSR 2-3) is also ligated. In this plasmid, the intact T4 lysozyme gene is present for the first time, but by fusion with the signal peptide gene and insertion in the lac Z gene. No unusually limited number of recombinant clones was observed here. The 3' untranslated part of the gene is removed by digestion with Dra I (in which the termination region is retained) and isolation of the large fragment. A Hind III linker is ligated directly to the (blunt end) Dra I cleavage site and after digestion to Hind III, the now liberated finished chimeric gene is cloned in the Hind III cleavage site of pUC 9 (pSR 2-10). Here again, the Hind III cleavage sites are converted to Sal I cleavage sites (pSR 2-12). Finally, the Sal I-flanked chimeric α -amylase-signal peptide-T4 lysozyme gene is cloned into the Sal I cloning site of the expression vector pAP 2034 in the correct orientation (pSR 2-4). Checking of the constructed plasmid is carried out with different restriction digestions (see Figure III/15).

```

5'  TGGCGTCC|ATGAATATATTGAAATGTTAC|GTAG  3'
    CCAACCGCAGG|TACTTATATAAACTTTACAATG|CATCCTAG
    Bgl II                      Sna BD    Bam HD

```

Figure III/13. Synthetic oligonucleotide for joining of the signal peptide and T4 lysozyme gene.

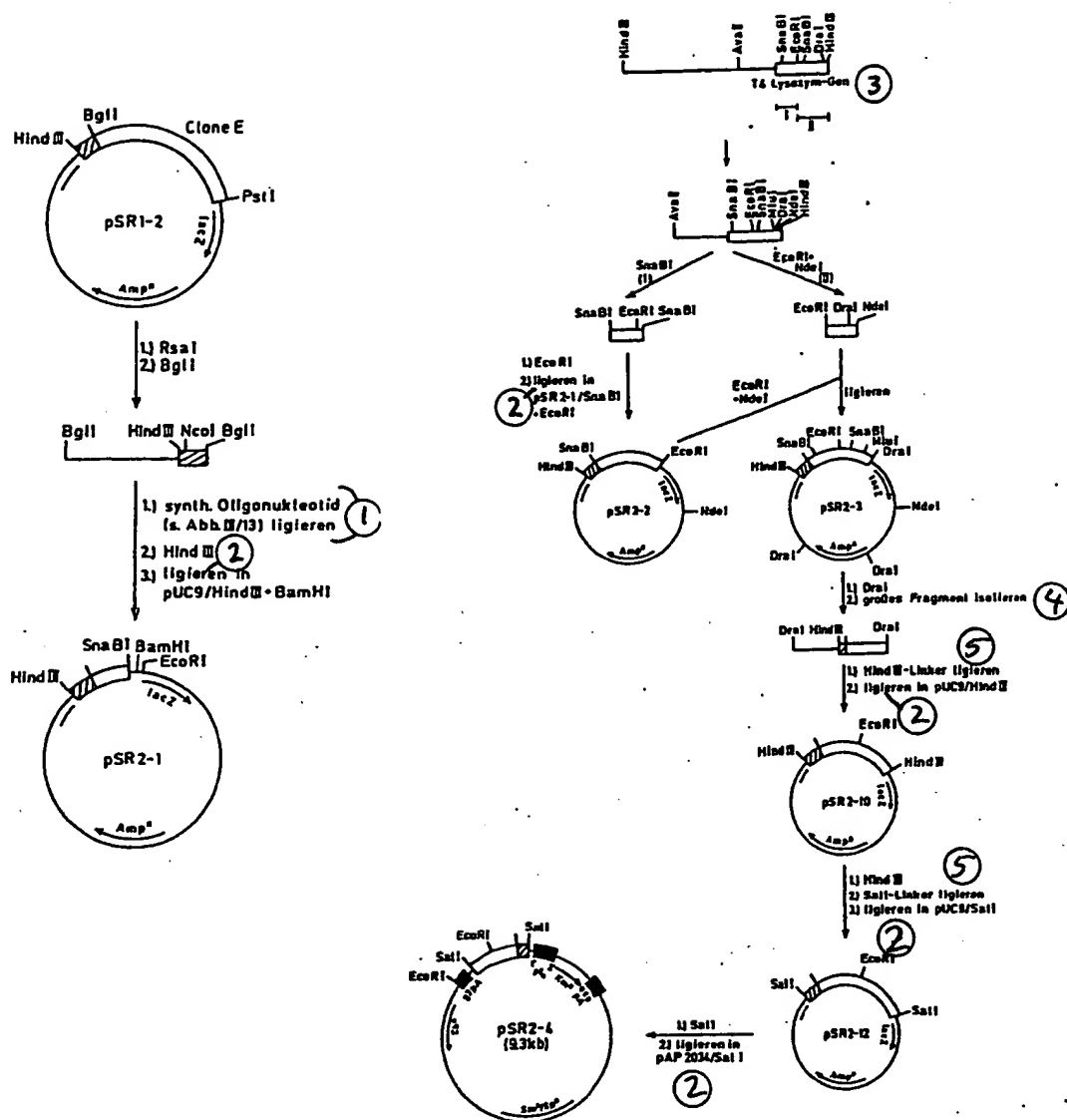


Figure III/14. Construction of a selection-expression vector for the chimeric α -amylase-signal peptide-T4 lysozyme gene under control of the T_R promoter:

- Isolation of the signal peptide gene
- Construction of the chimeric gene

Key:	1	Ligate synthetic oligonucleotide (see Figure III/13)
	2	Ligate
	3	T4 lysozyme gene
	4	Isolate large fragment
	5	Ligate linker

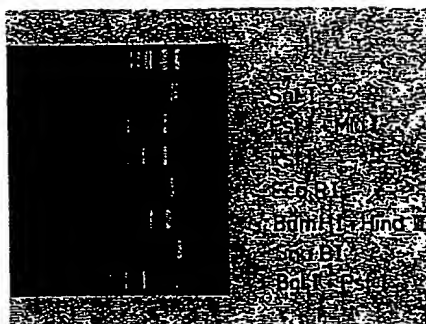


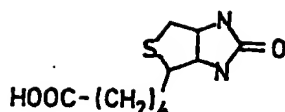
Figure III/15. Checking of plasmid pSR 2-4 by restriction mapping.

The plasmid DNA was digested with the stated enzymes and separated on a 0.7% TBE-agarose gel. λ -DNA digested with Pst I serves as length standard. Fragment sizes: Sal I: 8.6/0.67 kb, Pst I + Mlu I: 5.4/1.63/1.23/1.07 kb, Pst I: 5.4/2.3/1.63 kb, Eco RI: 8.75/0.58 kb, Bam HI + Hind III: 5.85/3.25 kb, Sna BI: 9.0/0.33 kb, Bgl I + Pst I: 2.3/1.6/1.05/0.8/0.42/0.3/0.2/0.1 kb.

The plasmid pSR 2-4 represents a selection-expression vector for the chimeric α -amylase-signal peptide T4 lysozyme gene. The fusion gene is situated under the control of the 1' position of the T_R double promoter and the g7pA region, the NPT II marker, and the selection gene in the 2' position, terminated by the ocs-pA region. At a size of 9.25 kb, it contains the carbenicillin, as well as streptomycin/spectinomycin resistance gene.

3.3 Nonradioactive labeling of DNA fragments for hybridizations by means of the biotin-streptavidin systems, development of a simple method

An attempt was made during this work to replace radioactive detection methods with modern enzymatic methods as much as possible. A large part of the Southern hybridization is carried out with plasma DNA or genomic bacterial DNA, in which the amounts of specific DNA being detected lie well above the detection limits. The now most commonly used system for sensitive detection is the biotin-streptavidin system. Biotin is a small, water-soluble vitamin (MW 244) and can be coupled easily via its carboxyl group to numerous biomolecules by different activation methods.



Streptavidin (from *Streptomyces avidinii*, MW 60 kDa) is a tetrameric protein with four highly specific binding sites ($K_D=10^{-15} \text{ M}^{-1}$) for biotin. In contrast to the initially employed avidin from chicken protein, streptavidin has more favorable physical properties (Chalet and Wolf, 1964; Green and Toms, 1970; Hoffmann et al., 1980), it causes significantly fewer nonspecific bindings and therefore leads to better results with equally strong affinity (Haeuptle et al., 1983). Because of its very high specificity, which is several orders of magnitude higher than in antigen-antibody reactions, only very short reaction times are required.

Several worked out methods exist for labeling of DNA probes:

1) Replacement of unlabeled nucleotides in the double strand by biotin-labeled nucleotides by nick translation according to the conventional method (Rigby et al., 1977; Maniatis et al., 1982). It has turned out that biotin 11-dUTP, in which the biotin group is bound by an 11-atom spacer to the C₅ position of the pyrimidine ring, is used as most favorable labeled nucleotide (Langer et al., 1981). About 30% of the thymidine residues are replaced by biotinylated uridine during the labeling reaction.

2) Terminal labeling by addition of bio-11-dUTP to the 3' hydroxyl group. The extent of labeling is limited by the number of available 3' hydroxyl terminal groups and both double and single strands are labeled, in which single-stranded biotinylated tails are obtained. The reaction by means of terminal-desoxynucleotide transferase is template-independent.

3) Bio-Bridge labeling system (Enzo)

This method is a further development of terminal labeling, in that here the 3'-poly dT tails are linked to the probe. This is denatured and hybridized on the filter-bound DNA. A biotin-modified poly dA molecule is then hybridized to the poly dT strand. The yield of biotin residues per DNA strand is higher here than in the two other methods.

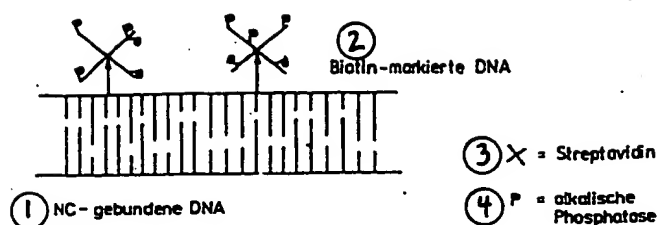
4) Oligolabeling (Feinburg and Vogelstein, 1983, 1984)

Kits for this labeling method have since been offered by Amersham (Multiprime labeling kit), Pharmacia (oligolabeling kit) and Boehringer Mannheim (Random primed labeling kit), but only the one from Boehringer Mannheim is suitable for biotin labeling. After denaturation of the double-stranded DNA, a new biotin-labeled strand is synthesized by means of a DNA single strand with DNA polymerase I (Klenow fragment). Synthetic hexanucleotides of all possible sequences serve as primer. dTTP and bio-11-dUTP in a ratio of 1:0.6 are used for the highest signal intensity.

Initially the labeling reactions were conducted routinely with nick translation, in which separation of the free nucleotide occurred by ethanol precipitation. After market introduction, this more costly method was replaced by oligolabeling, which makes isolation of the labeled DNA redundant and the entire reaction charge is used for hybridization. Because of the more difficult detectability of the DNA pellet (only possible by a complete detection cycle), this is an important advantage precisely during comparison with radioactive labeling. The oligolabeling method of Feinberg and Vogelstein in its modified form is therefore the method of choice for biotin labeling of DNA fragments. The hybridization conditions (Wahl et al., 1979) are only slightly changed by slight reduction of the melting point of DNA double strands that contain biotin-labeled double strands. The formamide concentration in the hybridization buffer is therefore reduced from 50 to 45% and the hybridization temperature remains unchanged at 42°C (Leary et al., 1983). Since proteins are used during detection, the nonspecific protein binding sites must be blocked in addition to the nonspecific DNA binding sites. A simple and cheap method is the use of 1% gelatin at room temperature and it yields no worse an effect than, say, 3% BSA, and even a lower background.

An important aspect in working out a detection method was universal applicability of the principle and the required buffers, for example, also for protein analyses by Western blotting, starting from the commonly present biotin group.

Tris was chosen as the buffer substance (10 mM) and 150 mM NaCl used for protein stabilization. In order to reduce the nonspecific background, detergents in low concentrations are used. Recently, it has been found that Tween 20 is superior to the previously predominantly used Triton X-100 and even can be used alone for blocking (Chan et al., 1985). Here 0.05% Tween 20 serves the aforementioned purpose. This buffer is known as TBST. For blocking, incubation at room temperature is carried out for 1 h in a solution of 1% gelatin (Bio-Rad) in TBST. With this TBST buffer, all possible additional detection steps can be carried out very well with antibodies or streptavidin. For detection of biotinylated DNA, only the complex with streptavidin to which a marker enzyme is coupled need be formed. After the previously commonly used marker enzyme peroxidase, after β -galactosidase and acid phosphatase, alkaline phosphatase is now increasingly gaining acceptance as the most sensitive enzyme. Nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) serve as substrates. By cleavage of the phosphate residue, the formation of a blue precipitating dye is initiated. For additional studies on the biotin streptavidin system, see 3.16. The streptavidin-alkaline phosphatase conjugate from the BluGene Kit, available from BRL, proved to be very well suited. An additional increase in the sensitivity of this system appears possible by optimizing the conjugate and by utilizing the snowball effect of the four biotin binding sites of streptavidin.



- Key:
- 1 NC-bound DNA
 - 2 Biotin-labeled DNA
 - 3 X = streptavidin
 - 4 P = alkaline phosphatase

Sensitivity of the system:

A sensitivity of about 500 fg homologous plasmid DNA could be achieved both with nick translation and oligolabeling, and in one concentration series during radioactive labeling, a limit of 1-0.5 pg with three days of exposure was achieved by oligolabeling. The detection limits for radioactive and biotin labeling are therefore comparable. As an additional time saving, as recently tested, UV crosslinking of DNA and RNA on Immobilon membranes for biotin labeling is also possible and in one DNA concentration series, a higher intensity of the signal was found during direct comparison for UV treatment (4 min on a 302-nm transilluminator with a moist filter) versus 2 h of baking at 80°C in a vacuum furnace. With viral RNA, a Northern hybridization was successfully conducted according to this method.

3.4 Integration of antibody construct in *Agrobacterium tumefaciens* and checking of the transformants by Southern hybridization

The previously described synthetically produced genes from plant regulation elements (promoter, polyadenylation sequence) and coding region for the desired proteins are integrated in the *Nicotiana tabacum* plants by means of the plant transformation technique developed a few years ago with *Agrobacterium tumefaciens* Ti plasmids (Matze and Chilton, 1981, Leemans et al., 1982; Zambryski et al., 1983; De Block et al., 1984; Horsch et al., 1984, 1985). The mobilization system of Van Haute et al. (1983) with the helper plasmids R64drd11 and pGJ 28, as well as the manipulated Ti plasmid pGV 3850 (Zambryski et al., 1983). This vector contains a pBR 322 fraction that is also contained in the selection-expression plasmid pAP 2034 (Vellen

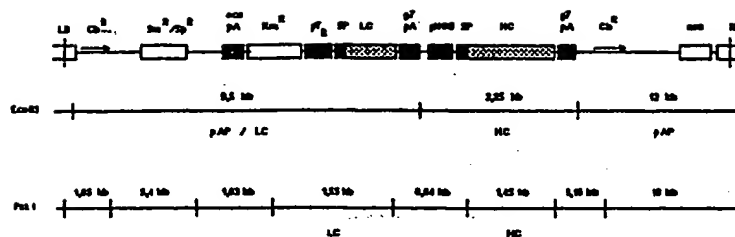
and Schell, 1986) so that the recombinant constructions can be introduced by homologous recombination to the Ti plasmid. The conjugation method is described in 2.5.3. During simultaneous selection, 10-100 colonies should grow per charge that are transformed with high probability.

Some selected clones are subjected to Southern hybridization analysis for intact incorporation of the foreign DNA. For this purpose, as described under 2.7.2, the total DNA from the agrobacteria is isolated in a miniprep. Aliquots are digested within different restriction enzymes and transferred to a membrane after agarose gel electrophoresis (see 2.7.4). Hybridization and detection with biotin-labeled probes occurs as described and discussed in 2.7.5 and 3.3. For the plasmid pSR 4-3, integrated in pGV 3850, the agrobacterial DNA is digested with Eco RI and Pst I. The 0.8-kb Bam HI fragment from pSR 1-10 (LC probe) serves as a hybridization probe for the light chain, the 1.55-kb Pst I fragment from pSR 3-10 (HC probe) for the heavy chain gene and the entire plasmid pAP 2034 (pAP probe) for the vector. The hybridizing fragments are graphically shown in Figure III/16, and the experimental result agrees with the expected one.

a)



b)



	Probe		
	pAP	LC	HC
<u>Eco RI</u>			
12.0 kb	X		
9.5 kb	X	X	
2.25 kb			X
<u>Pst I</u>			
1.45 kb		X	
1.45 kb			X

Figure III/16: Southern hybridization analysis of genomic DNA of pSR 4-3 transformed agrobacteria C 58 C1*.

a) Hybridization pattern with biotin labeling

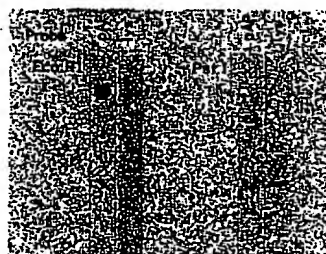
b) Graphic depiction of hybridizing reaction fragments

* [In the original document, commas in numbers represent decimals.]

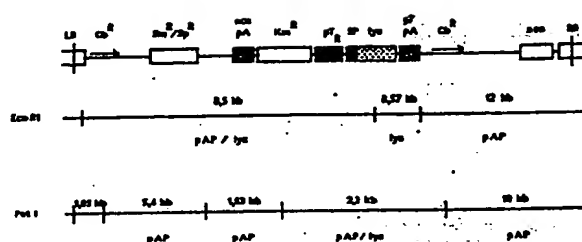
3.5 Integration of lysozyme construct in *Agrobacterium tumefaciens* and checking of the transformants by Southern hybridization

The same method as described in 3.4 is used for construction of pSR 2-4. The 0.67-kb Hind III fragment from pSR 2-10 (lys probe) is used for the hybridizations for the lysozyme genes and the entire plasmid pAP 2034 (pAP probe) for the vector. Figure III/17 shows the hybridizing fragments graphically and the experimental results agree with the expected ones.

a)



b)



	Probe	
	pAP	lys
Eco RI		
12 kb	X	
8.5 kb	X	X
0.57 kb		X
Pst I		
10 kb	X	
5.4 kb	X	
2.3 kb	X	X
1.63 kb	X	

Figure III/17: Southern hybridization analysis of genomic DNA of pSR 2-4-transformed agrobacteria C 58 C1.

a) Hybridization pattern with biotin labeling.

b) Graphic depiction of hybridizing reaction fragments.

3.6 Expression analysis of pT_R and pNOS

After some (RNA and protein) analyses had been conducted without result, it turned out that the T_R double promoter is obviously not expressed constitutively. Instead a nonuniform expression in the plants and inducibility by plant hormones was assumed (A. Szalay, C. Koncz, personal communication). Thereupon a detailed expression analysis of the promoter was undertaken in order to work out conditions under which an attempt could be made to successfully introduce proteins into the transformed plants. For this purpose, the NPT II gene under the control of the 2' position of the double promoter was used in the constructions of pSR 2-4 and pSR 4-3 and the activity measured by the NPT test (see 2.9.2) in the transformed plant materials.

It must be stated as a conclusion from all the conducted NPT tests that neither expression in the fresh state of the plants nor different inductions are absolutely reproducible and instead a large variability was found between the different transformants and also in proliferated plants that had already been analyzed once.

Expression analysis (conducted on sterile culture tissue) in intact fresh plants clearly shows increasing activity of the T_R double promoter from the tip to the base of the plant. The lowest expression is found in the highest leaf, directly followed by the hollow part of the stem. The activity in the leaves only increases slightly downward in most transformants and in some plants, however, a relatively high expression is already present in the lowermost leaf. Normally the highest and lowest leaf differ by a factor of about 1.5-2. It is then found that differences also occur in this leaf itself: the edge regions are more strongly active than the inner regions. Greater differences in activity are found in the stem, where a 25- to 40-fold higher activity is found from the top down. In the normal case, the roots exhibit the highest activity (about 45-fold relative to the highest leaf) and even here no uniform distribution could be established, but a weakening toward the root tip, which, however, is still about 15 times stronger than the highest leaf. As a conclusion to all these experiments, a region of greatest activity in the vicinity of the base of the plants at the transition from the stem to the roots therefore emerges. The expression rate in the fresh uppermost leaf is discussed below as a basic state of the T_R double promoter. A differential expression of this promoter in fresh plant tissue is therefore clearly demonstrated, but one cannot speak of tissue-specific expression, since the boundaries are fluid and variable. An expression gradient is involved instead.

These data suggest performing a study on the inducibility of the promoter. Knowledge of the assumption of hormone regulation (A. Szalay, C. Koncz, personal communication) was helpful for this. With this as point of departure, different hormone concentrations and ratios were tested and the most appropriate proved to be 2,4-D and kinetin in a ratio of 1:1 in a 10⁻⁵M concentration. The ratios 100:1, 10:1, 1:10 and 1:100 (in which 1 always means 10⁻⁵M) were also

tested, but somewhat weaker induction was found. For this purpose, cut off, whole leaves were incubated in liquid M + S medium for 5 days with the corresponding additive. Addition of hormones in this system is essential for induction. On the other hand, in a system for induction of stem pieces, the same induction is obtained both with and without addition of hormones. In this experiment the stem is divided into about 1-1.5 cm long pieces and then the length cut and the cutting surface placed on solid M + S medium and also incubated for 5 days. The difference between the two methods lies in the ratio of wound surface to total volume of tissue; in the stem a relatively large wound surface is present but in the cut-off leaves, only a minimal wound surface is present. In order to test whether merely gravity-related collection of hormones at the bottom of the segment is involved in this case (which corresponds to the cut surface), a special experiment was conducted. For this purpose two halves of a segment were rinsed in a fixed position with the cut surface up or down using liquid M + S medium and incubated for 5 days. Two tissue parts (upper and lower halves) were then cut out and analyzed. In both cases, the quarter with the cut surface exhibited the greater activity, which means that a gravity-related effect is not involved, but rather a wound-dependent one. This finding is important for the stem induction system, since normal transport of hormones is disturbed by cutting and the change in direction. An attempt was made in leaves to obtain a stronger induction by enlarging the wound surface. For this purpose, leaves were initially cut into strips and incubated in the described manner in liquid M + S medium for 5 days. A positive effect was found. Later the transition to solid M + S medium occurred and the leaves were cut into small pieces, as in a leaf disk test. These changes caused adaptation of the two (leaf and stem) systems to each other. In both systems, the same induction could not be achieved. If previous induction of leaves was only possible by addition of hormones, now it occurred merely by wounding and to a greater extent (two to four times higher).

The T_R double promoter is inducible in stem pieces in vitro about 70 to 170 times, in leaf pieces about 25 to 150 times in comparison with the base state. The CaMV 35S promoter analyzed at the same time in calli as a reference (provided by R. Topfer) is always about 1000 times stronger in the NPT test than the base state of the T_R double promoter.

Kinetic studies were conducted both for stems and for leaves, with and without hormones. The increase in activity of the promoter begins during wounding without hormone addition in both cases about 1 day after the beginning of induction and rises slowly and from the third day quickly. The investigation was conducted up to 6 days, the highest induction being measured on the sixth day. According to this pattern a similar experiment was conducted by M. Wassenegger (Cologne, personal communication) with a luciferase test over a period of 11 days. The maximum expression was also on the fifth to sixth day followed by a uniform decline to roughly the base state on the 11th day. During the addition of hormones, the same high activity as

after 5 days was measured in the stem already after 10 h, whereas no change was noted in the intact leaves. No difference in inducibility of the upper or lower leaves of a plant at the same time could be found in different experiments, but high variability was found between different plants and proliferated examples of the same plant.

The natural optimal expression sites of the T_R double promoter are found in calli, which develop after infection of the plant with *Agrobacterium tumefaciens*. Good expression should therefore also be expected in sterile culture calli. This was already demonstrated by Velten et al. (1984) and confirmed by M. Prols and R. Topfer (Cologne, personal communication). In the transformed tissue here, the strongest expression with a margin was found both in fresh and induced tissue in the calli and only one plant reached this activity in the leaf and stem tissue. The activity in the callus is comparable to that of the CaMV 35S promoter, which is generally considered constitutive and the strongest now used promoter. But differences are also found here. After repeated tests, it must be stated that expression in soft and unorganized calli is stronger than in hard, white or green calli. Calli that formed on the lower stem of the single lysozyme plant L2 exhibited only about 20% of the activity of the 35S calli and therefore corresponded more to the hard-green callus.

It was known from M. Prols (1966, personal communication) and R. Topfer (personal communication) that an equally strong expression as in calli was also observed in carrot protoplasts. These results could not be reproduced in tobacco mesophyll protoplasts. Generally only a relatively weak signal could be obtained there and, even under kinetic conditions, a relatively weak signal that diminished further by the fifth day was found after 2 days. The use of different hormone concentrations and ratios had no effect in protoplasts. The plasmid pAT 2034 (Velten and Schnell, 1986) was used for transient expression in tobacco protoplasts (contains 1'-NPT). In other groups, good expression of the double promoter could sometimes be achieved with different plasmids both after 1 day and also after 3 days (I. Moore, R. Masterson, D. Wing, personal communication), and in this case pUC derivatives function more poorly than other plasmids (R. Masterson, personal communication). During comparison of transient expression of both plasmid pSR 2-4 and pSR 4-3, it was found that the lysozyme construction after 1 day of incubation achieved about 1/7 of the activity of the 35S promoter, but antibody construction achieved only about 1/35; obviously the presence of T4 lysozyme influences the activity of the promoter positively. This will be taken up further later.

The possible high expression in calli and protoplasts appears logical, since in both cases tissues that are in a strongly activated state are involved. On the other hand, unfavorable conditions prevail in intact plants. There the 35S promoter is 1000 to 20 times stronger, whereas after induction of the double promoter, it is still only 20 to 5 times stronger. It is still only

equally strong in unorganized calli. This means that the double promoter in its highest induced state is an equally good (active) promoter as the CaMV 35S promoter.

Since a clear influence of promoter strength by wounding occurred in these experiments, a train of thought to also test acetosyringone, which is known to induce the vir gene of the Ti plasmid and is formed by wounded plant cells (see Introduction) for promoter induction, arose. However, no clear picture could be obtained for this substance. Sometimes quite good induction results were obtained, but sometimes no difference was found relative to simple M + S (wound) induction with whole leaves. This substance might be directly related to wound reaction, but this direction was not pursued further.

In a large NPT test all available transformed plants were tested for strength of expression of the 2' NPT gene. For this purpose, the lowermost stem piece was freshly frozen and the next higher piece used for induction. It was divided into two equal halves through a longitudinal cut, one of which was induced on solid medium without hormone addition and the other with 1:1 hormones for 5 days. It turned out that some plants are only poorly inducible or not at all, but most have a roughly equally strong base state and are clearly inducible. The picture of optimal induction, however, is ambiguous, in one case induction only being stronger by wounding, in others by addition of hormone. The induction factors are also different from plant to plant. It must be stated in conclusion that the developed inducibility does not yield a clear, readily reproducible picture but instead a very variable one that suggests a complex nature of regulation.

To clarify whether expression of the double promoter might occur in agrobacteria, which could lead to their damage in the case of pSR 2-4 by T4 lysozyme and thus prevent transfer of plasmid DNA to wounded plants, the activity was also investigated in bacteria. Gelvin et al. (1981, 1985) and DiRita and Gelvin (1987) already found weak activity in agrobacteria and in *E. coli* for a similar promoter from another Ti plasmid. A significant activity in *Agrobacterium tumefaciens* and *E. coli* was actually also found for this promoter, but no inducibility by hormones or acetosyringone could be found. Since, however, the mechanism of promoter regulation is unclarified, it cannot be ruled out that induction also occurs in the agrobacteria due to some substance formed by the wounded plant tissue.

Later a transfer of the induction conditions from sterile culture plant to greenhouse plants had to occur in order to have the required amounts of plant material for protein analyses available. Since no induction can occur here on an artificial nutrient medium, the conditions had to be adjusted. For induction of leaves in greenhouse plants, the leaf surface is roughened with a toothbrush and slightly wounded, whereupon a solution of 2,4-D and kinetin (both 10^{-5} M/1:1) is sprayed on the wounded leaves. Harvesting was initially carried out after 5 days, later after 2 to 3 days, since after a few days wounding of the leaves leads to partial dying of the tissue. After 2 days most of the leaf tissue is still intact. As in the sterile culture, plant L2 reproducibly

exhibits a very high base level (the highest of all plants), but is not inducible. Plant A10 behaves differently, with relatively low expression in the uninfluenced state, but is insignificantly inducible. A large jump occurs from the period before induction to the first day after it and up to the fifth day only a slight additional rise follows. For comparison, the 35S promoter in A10 is about 15 times stronger after induction of the double promoter, in L2 only about equally strong in the leaves! Induction therefore also functions in active plants, in which it was not investigated whether wounding or the use of plant hormones is the decisive factor. The objective of the overall series of studies to find optimal conditions for expression of the T_R double promoter for protein analyses and estimate the strength of expression is therefore achieved.

In order to ensure that the 1' promoter is also regulated in the same way, only two methods could be resorted to: in the first place, analysis of transformed plants that contain the 1' NPT gene in a different environment (furnished by G. Coupland, Cologne) and, on the other hand, transient expression of a plasmid analogous to pAP 2034 (pSR 6-1). This plasmid was constructed by deletion of the Sal I fragment, which contains the bacterial kanamycin resistance gene from Tn903, from pAK 1003 (Velten et al., 1984) and contains the 1' NPT gene in the same environment as pAP 2034 (Velten and Schell, 1986). During transient expression of both plasmids, no difference was found in the promoter strain. The 1' NPT plant could only be tested in a very young state and exhibited the highest expression in the lowermost leaf, but not in the stem or root. A significance of this result must first be demonstrated by repeating the experiment. The 1' promoter, however, reacted to the induction conditions in the same way as the 2' promoter. It is also 35 to 70 times inducible in both induction systems (leaf pieces and stem segments) within 5 to 6 days, both by wounding and with hormones. An equivalent induction with the same base expression is therefore reliably assumed.

The same analyses were also conducted with the NOS promoter in order to test whether an analogous controllability is present despite all previous publications. Since both promoters control genes with similar function (opine-synthase genes) a similar regulation would not be unfounded. No such studies have yet been published either. The activity analysis in fresh tissue also produces differential expression for the NOS promoter, which increases from the top down. In both tested plants, a three-fold higher activity is uniformly present in the lower leaf, stem and roots than in the upper leaf and stem. The inductions are about six-fold for wounding and 20-fold for hormone addition (1:1) in leaf pieces and in both cases about eight-fold for the stem pieces. The qualitative effects are also the same and are only quantitatively lower. The base states of both promoters lie on the same order of magnitude and the inductions can be higher in the double promoter. After culturing of calli, these were also compared, and the same conditions predominated here: hard, green or white calli expressed more strongly than a soft, unorganized

one and in such a callus the expression is about four times as high as in the lowermost leaf, i.e., about 8 times above the base state.

The following picture therefore emerges: pNOS and pT_R are qualitatively both regulated according to the same principles, but have quantitative differences in the factors. In the base states, they are comparably strong, both are wound- and hormone-inducible, but the double promoter to a stronger degree. The same applies to expression in calli. However, the double promoter has major advantages, because it can reach the strength of the 35S promoter (as a reference in the callus), which means a fragment of total protein of about 0.1% (Sanders et al., 1987). Contrary to the measurements of Sanders et al., the factor between NOS and 35S promoter in the NPT test is not only 110, but even about 1000 in the base state. Comparisons at the RNA level were conducted by Sanders et al. (1987) and Harpster et al. (1988) and the results of both studies are similar.

For protein analysis for the light and heavy chains of the antibody, the results proved to be favorable, since the activities of both promoters correlate in order of magnitude. On the one hand, a roughly comparable amount of light and heavy chain is present in the induced leaf tissue, which should facilitate assembling and on the other hand, however, under optimized conditions, if necessary, a very high percentage of light chain which cannot be glycosylated can be produced. Major differences between both protein amounts would distort the primary picture.

In the F₁ generation of the L2 plants, all NPT-positive examples also show the high activity known from the parent plants.

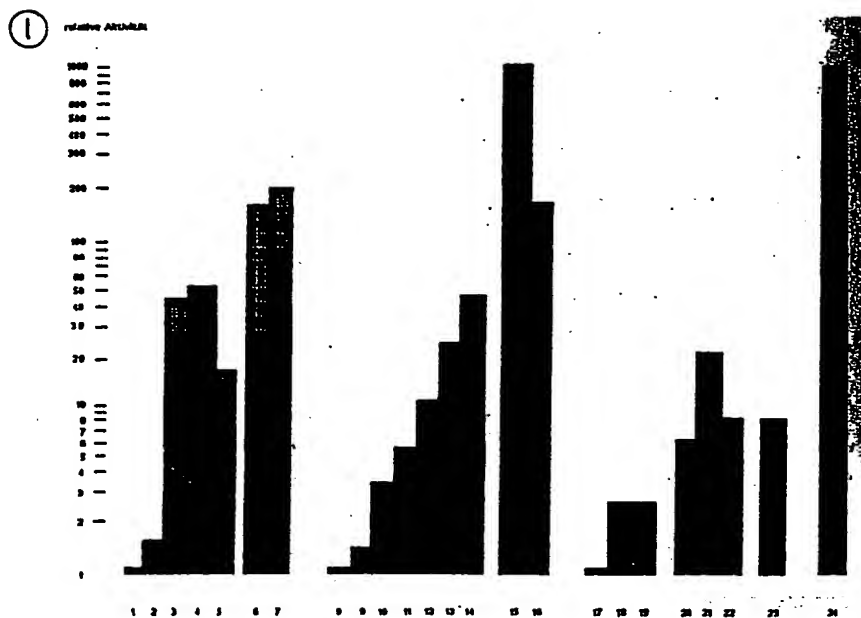


Figure III/18: Average values of differential expression in different plant tissues and induction in leaves and stems in sterile culture plants (relative activity plotted logarithmically, base expression in uppermost leaf = 1).

1-16: pTR-2'-NPT II (1-5 fresh), 1: uppermost leaf, 2: lowermost leaf, 3: lowermost stem piece, 4: middle root, 5: root tip; (6-14: induced with wounding and 1:1 hormones (see text)), 6: leaf, 6 days, 7: stem, 6 days; 8-14: kinetics (leaf) 5 h, 14 h, 1 day, 2 days, 3 days, 5 days, 6 days, 15: callus, 16: tobacco mesophyll protoplasts, 17-23: pNOS-NPT II (17-19: fresh), 17: uppermost leaf, 18: lowermost leaf, 19: lowermost end piece, (20-22: induced, 5 days), 20: leaf by wounding, 21: leaf with 1:1 hormones, 22: stem by wounding or with 1:1 hormones, 23: callus, 24: pCaMV 35S-NPT II in callus as reference. The brightly shaded surfaces: maximum values, additional explanations in text.

Key: 1 Relative activity

3.7 Expression of lysozyme construction pSR 2-4 in *Acetabularia mediterranea*

The correct expression of genetic engineering construction can be checked in a system of microinjection of DNA established with the one-celled green alga *Acetabularia*. *Acetabularia* is an appropriate system for investigating expression of heterologous genetic information (Neuhaus et al., 1983, 1984, 1986; Langridge et al., 1985; Berger et al., 1987; Schweiger and Neuhaus,

1987). About 10^7 supercoiled DNA molecules are microinjected with a concentration of 1 mg/mL into isolated cell nuclei of *Acetabularia* (about 10 μ m diameter). In parallel, denucleated *Acetabularia* are isolated into which the microinjected nuclei are reimplanted by fusion. The overall treatment leaves the stability and morphogenetic activity of the nucleus uninfluenced. The manipulated *Acetabularia* are tested in molar medium (Primke et al., 1978) for capacity for expression of the heterologous genetic information.

Detection of the expressed protein occurs by indirect immunofluorescence. For this purpose, the transformed *Acetabularia* are spread on a microscope slide, dried and fixed. They are then incubated with the specific antibody and then with fluorescein-coupled second antibody. The immune reaction is made visible by the fluorescence.

The described investigation method was applied to lysozyme construct pSR 2-4. The polyclonal rabbit anti-T4 lysozyme antibody described in 2.8 was used as a specific antibody. After the immune reaction, a specific immunofluorescence could be detected in the transformed *Acetabularia*. Controls were conducted. The correctness of the construction in plasmid pSR 2-4 was thus demonstrated and it was found that T4 lysozymes could be produced in the plant cells. A clear indication of secretion of T4 lysozyme could not be obtained. A further conclusion is the functional capability of the T_R double promoter under the culture conditions of *Acetabularia mediterranea*.

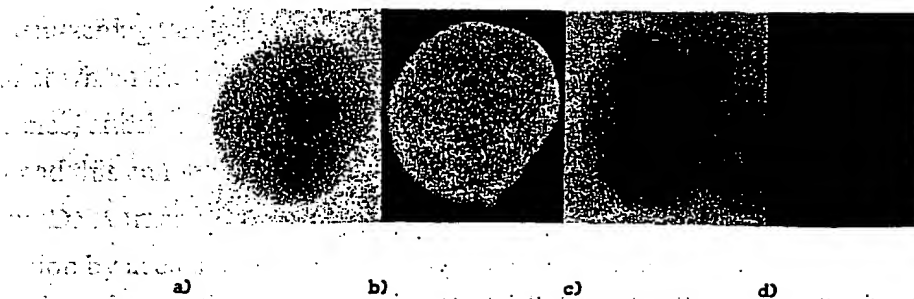


Figure III/19: Detection of expression of T4 lysozyme from pSR 2-4 in *Acetabularia mediterranea* by indirect immunofluorescence.

- a) Light field recording of object
- b) Dark field recording of the UV light with specific fluorescence for T4 lysozymes
- c) + d) Negative control of pBR 322-DNA: light field (c) and dark field recording (d).

3.8 Transient expression in tobacco protoplasts

Transient expression in tobacco protoplasts can produce quite good activities of the double promoter, but only very small amounts of total protein can be isolated from a charge of

the main protoplasts. Repeated attempts to directly detect the light chain of B 1-8 and for T4 lysozyme from the crude extract of tobacco mesophyll protoplasts were unsuccessful. An increase in the number of protoplasts to orders of magnitude so that a positive single would be expected is connected with unacceptably high expenses both in work and cost. For this reason, this approach was no longer pursued. The transient expression in protoplasts is otherwise suitable only for in situ detection, for example, by immunogold labeling.

3.9 Transformation of *Nicotiana tabacum* W38 with antibody construct and transformation analysis

The chimeric genes integrated in the Ti plasmid pGV 3850 are transferred to *Nicotiana tabacum* W38 by agrobacterial infection (Zambryski et al., 1983). This occurs by means of the leaf disk test (Horsch et al., 1985) in modified and simplified form (see 2.6.2). In the simplest and fastest method, separate callus induction is no longer carried out but it is placed right on the sprout induction medium with 0.5 mg/L BAP and 0.1 mg/L NAA.

Two leaf disk tests were set up with the agrobacterial described in 3.5, one of them according to this method and another with coinfection (see 2.6.2). Half were cultured with and without kanamycin selection (50 mg/L). Calli that begin sprout formation after some time are obtained during coinfection. This method requires a longer overall time. During direct sprout induction, small calli that produce sprouts are sometimes obtained and sometimes small sprouts within 4 to 6 weeks directly on the leaf surface. These sprouts are cut off and selected with 100 mg/L kanamycin. From the coinfection charge, five transformed plants for the antibody construct and according to the other method 15 transformed plants were regenerated. Transformation analysis is done by the nopaline and NPT test by means of the two marker genes. The nopaline gene is located in the T-DNA of the Ti plasmid next to the right border; the NPT gene is under the control of the 2' position of the T_R double promoter (from plasmid pAP 2034, Velten and Schell, 1986) and is subjected to the tissue-specific expression analyzed in 3.6. Nevertheless, the activity of the promoter is sufficient to generate resistance against 100 mg/L kanamycin in the buds. A first index of transformation is good growth under kanamycin selection and especially rooting under these conditions. At the time of the nopaline test (Otten and Schilperoort, 1978) for detection of the presence and activity of the nopaline gene most buds were already rooted under kanamycin selection. Clear yellow spots for nopaline could be obtained in paper electrophoresis (see 2.9.1).

Nopaline-positive plants were subjected to an NPT test (Reiss et al., 1984; Schreier et al., 1985) (see 2.9.2) in order to also detect the presence and activity of the second marker gene. Part of the nopaline-positive buds did not survive continuing selection or exhibited no signal in the NPT test. There is a general explanation for this: in recent years strong variability in expression

of several regenerated plants from a transformation charge has always been demonstrated by different groups, even from a callus, which is attributed to a "position effect" (Jones et al., 1985). It is very often observed that only one of the two present marker genes (NOS and NPT) is expressed or that deletions occur (Budar et al., 1986; Czernilofsky et al., 1986; Uchimiya et al., 1986; Sanders et al., 1987). It is also known that the degree of methylation can influence the activity of the nopaline synthase gene (Hepburn et al., 1983). On the other hand, the now known differentiation between different parts of the plant with reference promoter activity must be considered. Initially, young leaves from the tip of the plant are always used for the NPT test, which, however, have the lowest or almost no activity, as could be demonstrated in this study.

For further analyses, 20 transformed plants, both nopaline- and NPT-positive, were used. No assertion can be made concerning the relative activity of the introduced genes by means of the NPT test because of the nonconstitutive expression of the promoter, since leaf material is used. For this reason, a larger NPT test was conducted for comparison of all plants in which the same materials (fresh and induced) were compared with each other. Layer tests with proliferated material from the same plants, however, show that the induction conditions are in no case absolutely reducible and the activity of the promoter is therefore always subject to greater fluctuations and only a tendency for relative activity of the plants can be suggested.

The plants A4, A5 and A8 exhibit the highest expression in this test, in which A5 is by far the strongest and for additional analyses, plants A3 and A10 were also used. All the mentioned plants exhibit a single specific band for the light chain or heavy chain gene in Southern hybridizations (see 3.11), i.e., no rearrangement occurred and the genes are therefore incorporated properly.

3.10 Transformation of *Nicotiana tabacum* with lysozyme construct and transformation analysis

For the lysozyme construct (see 3.5), a leaf disk test (*Nicotiana tabacum* W38) with the usual hormone concentrations (1 mg/L NAA + 0.1 mg/L BAP for callus induction and 1 mg/L BAP for sprout induction) (Zambryski et al., 1983; Kaulen, 1986) was initially carried out. Under these conditions, however, no calli could be proliferated. Thereupon, an additional five leaf disk tests were conducted according to the methods explained in 3.9, among which was a coinfection from which only two transformants could be regenerated. One was nopaline-negative (see below/L2), the other NPT-negative. Additional transformants could not be determined either with or without selection. For this reason, the last leaf disk test was conducted with acetosyringone induction. As could recently be demonstrated, this substance formed from wounded plant tissue induces the var gene from the Ti plasmid of *A. tumefaciens* (Koukolikova-Nicola et al., 1985; Stachel et al., 1985, 1986a/b; Bolton et al., 1986; Wang et al.,

1987) and on this account causes infectiousness of the agrobacteria. It could be demonstrated for *Arabidopsis thaliana* that during treatment of the agrobacteria with 20 μ M acetosyringone (16 h, pH 5.6) before incubation the transformation frequency is significantly increased (Sheikholeslam et al., 1987). The following concept was developed from this: the agrobacteria culture is induced for 16 h in LB (pH 5.6) under antibiotic selection of 20 μ M acetosyringone. The leaves are treated for two days with 100 μ L of this culture in L + S medium (pH 5.6) and 20 μ M acetosyringone. After a week selection is conducted. This approach also led to no result. An explanation is obtained from the subsequently performed activity analysis of the T_R double promoter. In the first place, the tissue-specific expression can have an effect on selection capacity, which is less likely, however, because of the number of transformants obtained for the antibody construction in this extent. A key idea can be derived from the inducibility of the promoter by acetosyringone. As described under 3.6, it could be demonstrated that the double promoter under normal culture conditions is active in agrobacteria. A weak activity of the T_R -2' gene of another Ti plasmid has already been described in bacteria in the literature (Gelvin et al., 1981, 1985; DiRita and Gelvin, 1987). However, production of lysozyme cannot be sufficient for lysis of bacteria; since normal growth was observed. In addition, the inducibility of the promoter was detected by wounding the plant tissue. It is therefore assumed that lysozyme production in the agrobacteria is activated at the moment when these are cultivated together with the wounded plant pieces in the disk test. In similar fashion to the natural action of T4 lysozyme (production in the attaching bacteria and lysis of the bacterial cell wall from the inside, Tsugita, 1971). At the point at which the DNA transfer machinery of the Ti plasmid is set in motion by activation of the vir genes, attack of newly produced T4 lysozyme is assumed simultaneously on the bacterial cell wall and this can lead to an adverse effect on hampering of the function of these cells walls during DNA transfer, whose mechanism is not clarified, or to destruction of the bacteria. Induction by acetosyringone cannot be ruled out in conjunction with wound-inducibility so that acetosyringone addition during the leaf disk test could then even have negative consequences. Transformation of plant cells in this case must therefore be assumed as an exceptional phenomenon.

For this reason a transformation of *Nicotiana tabacum* SR1 (Maliga et al., 1973) was conducted by direct DNA transfer in the weak protoplast (see 2.6.4) in order to get around the problem of lysis effect of T4 lysozyme on agrobacteria during the transformation process (cooperation with R. Hain, Monheim).

The regenerated plant L2 was analyzed together with the antibody plants in the NPT test and the relative activity determined (see Figure III/20). It also turned out in later NPT tests of these plants both from the greenhouse and from the sterile culture that they have a very high base level of expression (the highest of all plants), but are not inducible. In the Southern hybridization

analysis, only the correct bands were found. The base level of expression increased with increasing age of the plant.

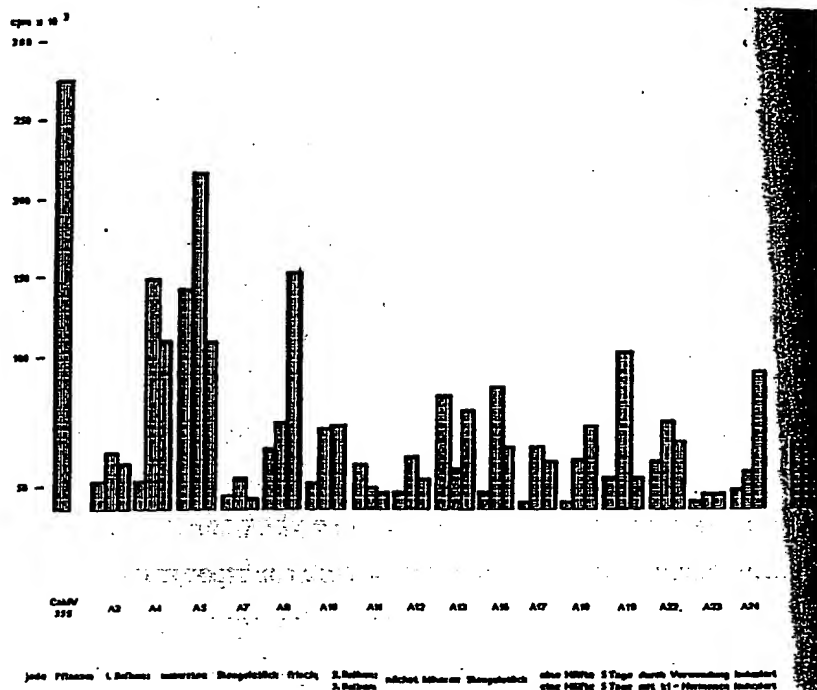


Figure III/20: NPT test in comparison of regenerated transformants. The lowermost stem piece in each plant was freshly analyzed and the next higher one induced in two halves for 5 days only with wounding on M + S medium or with hormone addition (1:12, 4-D and kinetin 10^{-5} M). The CaMV 35S-NPT II callus serves as comparison.

Key: 1 Lowermost stem piece, fresh
 2 Next higher stem piece, 1/2, 5 days induced by wounding
 3 Next higher stem piece, 1/2, 5 days induced with 1:1 hormones

3.11 Integration analysis of antibody plants by Southern hybridization

The isolated genomic plant DNA (see 2.10.1) is digested with restriction enzymes – here double digestion with Eco RI and Hind III – separated on 0.7% TBE agarose gel and transferred to a membrane filter by Southern blotting.

Hybridization occurs with a radioactively labeled specific probe (see 2.10.2) under adjusted washing conditions. The biotin labeling was repeatedly used and different washing conditions tested, but a satisfactory, reproducible result could not be achieved and routine

applicability would actually be expected of the sensitivity of the system. For this reason, radioactive labeling was resorted to in this case. The 0.9-kb Sal I fragment from pSR 1-16 (light chain gene with signal peptide (LC probe)) was used as the probe for the light chain and the 1.7-kb Hind III fragment from pSR 3-10 (heavy chain gene with signal peptide (HC probe)) was used for the heavy chain. As a result, a hybridizing band of 3.4 kb is expected for the light chain and one of 1.95 kb for the heavy chain.

Only plants that show an intact integration pattern according to DNA analysis were used for further analysis.

3.12 Integration analysis of lysozyme plants by Southern hybridization

The investigation was conducted according to the same method as in 3.11. The plant DNA in this case is only digested with Eco RI. The 0.67-kb Sal I fragment from pSR 2-12 (T4 lysozyme gene with signal peptide (lys probe)) served as the hybridization probe; two hybridizing bands of 8.7 kB and 0.57 kB size are expected. In the case of plant L2 these bands are precisely obtained, and for plant L1 roughly 9-kb and 2.4-kb bands are found. A rearrangement therefore appears to have occurred here and no cross-hybridizing bands are observed with W38 DNA (see Figure III/22).

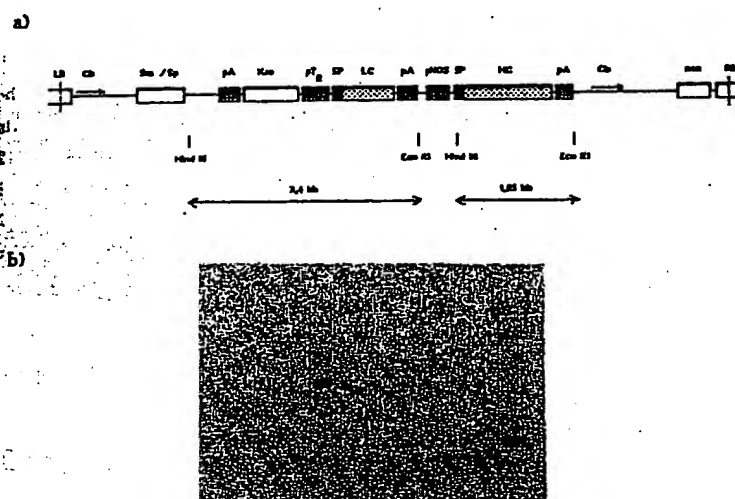


Figure III/21: Southern hybridization. Analysis of genomic DNA from transformed pSR4-3 tobacco plants.

a) Graphic depiction of restriction fragments

b) Typical hybridization pattern of an antibody plant

(For the HC probe a 2.4-kb band that cross-hybridizes with wild-type DNA was not used for evaluation (marked "x"))

Logemann, dissertation in preparation). In this case, a hybridizing band could be repeatedly found with plasma pSR 2-4 in the Northern blot).

3.14 Expression analysis of antibody plants by Western blotting

For the analysis of the transformed plant tissue for the presence of the sought proteins, immunological detection methods like Western blotting should be used. For identification of the individual chains of B 1-8 in the entire molecule, a number of different polyclonal and monoclonal antibodies are available (furnished by A. Radbruch):

Ls 136: monoclonal antibody that recognizes the light chain in undenatured form (also individually)

anti- λ , biotinylated: polyclonal antibody for detection of the light chain in native or denatured state

anti- μ Ig, biotinylated: polyclonal antibody used for detection of the heavy chain but also weakly recognizes the light chain

Ac 38: monoclonal antibody whose epitope is formed from both chains of B 1-8 and lies near the active center; it recognizes only intact B 1-8 with the correct tertiary and quaternary structure

Ac 146: monoclonal antibody that recognizes a structure in or in the immediate surroundings of the active center of B 1-8; possibly an isolated heavy chain can also be recognized in extremely limited frequency

NP/NIP: haptens for B 1-8, the monoclonal antibody B 1-8 was formed opposite the hapten NP, but exhibits heteroclytic properties and recognizes the related hapten NIP with greater affinity

(Reth, 1981, personal communication)

Due to the expected low relative percentage of sought protein in the total protein, sensitive detection methods had to be developed. For this purpose, a nonradioactive detection system, which is compatible with the system for DNA hybridization described in 2.7.5 and 3.3, was worked out for detection of proteins on nitrocellulose and Immobilon membranes.

As already mentioned there, alkaline phosphatase is the most sensitive currently known detection enzyme and is more sensitive by roughly a factor of 50 than the most widely used peroxidase. Our own studies also showed that in a direct comparison between alkaline phosphatase-coupled second antibody and biotin-labeled second antibody with alkaline phosphatase-coupled streptavidin, an increase in sensitivity by a factor of 15 additionally occurs. This combination permits the same sensitivity as the very demanding radioactive labeling with ^{125}I .

By using 1% gelatin as the blocking reagent in TBST buffer, a fairly low background coloring can be achieved. It is important here to work under mild conditions (as stated here) so that bound proteins are not washed out again from the membrane under the selected conditions. The presence of Tween 20 during the entire incubation and washing reduces the nonspecific signals additionally. By replacing nitrocellulose with the Millipore Immobilon membrane, the background colors attributed to mechanical damage are reduced to a minimum and contrast is also increased. The method described in 2.12.8 offers a highly sensitive analysis method for immunodetection under problematical handling. The absolute detection limits depends strongly on the corresponding proteins and the employed antibodies. For detection of the light chain from B 1-8 with a polyclonal biotinylated anti- λ antibody, a sensitivity of about 30 pg light chain can be achieved in the dot test. However, other combinations can produce much less sensitive detection limits.

In combination with the immunological part of Western blotting, a time and cost-saving new blotting method, semidry blotting (Kyhse-Andersen, 1984), was worked out and used. Transfer of the proteins from the polyacrylamide gel membrane occurs according to the same principle, but only with moistened filter paper between two carbon electrodes, and the time is shortened from at least 16 h to only 25 min. Because of this, large amounts of blotting paper are saved and during use of minigels, a Western blot could be completely conducted in one day.

Comparison of the maximal activity of the double promoter estimated in this work by promoter analysis (see 3.6), which is comparable to that of CaMV 35S promoter (about 0.1% of the total protein), makes it clear that only under optimal conditions (strongly active tissue) can a direct detection of the light chain be possible from the crude extract in the Western blot. Since the NOS promoter could not reach the activity of the double promoter even during good induction, a direct detection from the crude extract is not to be expected under any circumstances in this case.

A method therefore had to be developed that permits the sought protein to be enriched from the crude extract before Western blot or preferably to be isolated and concentrated to detectable concentrations. An affinity chromatographic method offers the best possibility of conducting a clean isolation with simple aids. Most of the other available methods permit only partial enrichment of the protein being analyzed or require greater apparatus expense. In addition, the method must be suitable for investigating larger numbers of samples with tolerable expense.

The series of different first antibodies for immune reaction with the sought proteins can be easily immobilized on CNBr-activated Sepharose 4B (see 2.8.3). An appropriate matrix is then available for their binding. The plant material is homogenized in liquid nitrogen in a mortar with a pestle and transferred to a centrifuge tube. The powder is mixed there with extraction

buffer and PVPP and extracted for 15 min on ice. The polyvinylpyrrolidone (PVPP) binds phenolic compounds without also absorbing the sought proteins; other materials, like Dowex 1X2-ion exchange resin, which was also tested, could not be used because of binding of these proteins. For protection against proteases which are liberated in larger amounts from the vacuoles, protease inhibitors must be added. A mixture of PMSF, leupeptin (against serine proteases), pepstatin (against acid proteases) and EDTA (against metalloproteases) should convey adequate protection. 5 mM ascorbic acid is also added for protection against oxidizing substances in the crude extract. 100 mM PBS is used as buffer for the soluble extract. If membrane-bound or proteins enclosed in the membranes are also to be liberated, a detergent must be added. For this purpose, 1% Triton can be used or preferably 10 mM CHAPS. After centrifuging of the extract to 10,000 rpm, the supernatant is diluted to 1:5 and rolled on a Denley roller at 4°C for 30 to 60 min. During this time part of the dissolved oligosaccharides precipitate from the plant extract. This precipitate is centrifuged and protein determination is carried out in order to have equal total amounts of protein in the experiment and control. After overnight incubation of the corresponding affinity gel at 4°C on the Denley roller, the bound protein is eluted in an Eppendorf flask under appropriate conditions and concentrated by ultrafiltration. After addition of SDS probe buffer, the finished probe is applied to Western gel and analyzed.

During immunodetection of the filter-bound proteins, nonspecific signals can occur with the first and second antibodies from overcharging and saturation phenomena; streptavidin gave no nonspecific signals in one test. These can be reduced or entirely eliminated if, on the one hand, the filters are washed after incubation with the antibodies with 1M salt solution (TBST/1M NaCl) and on the other hand, and much more effectively, the nonspecific binding sites of the antibodies, by preincubation of the antibody solution, are saturated with a wild type plant extract for 10 min at room temperature. Moreover, 0.1% Triton X-100 and 0.1% Tween 20 can additionally be used as the detergent.

A simple rapid analysis is tissue printing (Cassab and Varner, 1987). By pressing out tissue parts on a membrane, all the present proteins are essentially bound. There are no problems with this as in the production of extract. By using this method for calli obtained from transformed plants (A5, A10), positive signals for all the sought proteins can be found (light chain, heavy chain, and aggregated antibodies B 1-8), in which case the chlorophyll staining was clearly covered by the color development of alkaline phosphatase. By saturation of the employed antibodies as wild type extract, no cross reactions develop for the control spot and the green color of the chlorophyll remained visible. Positive signals for the light chain can be seen both with a polyclonal anti- λ and with monoclonal Ls 136 (anti- λ) antibodies. With the polyclonal anti- μ Ig antibodies, the presence of the heavy chain could also be detected. By detection with the

Ac 38 antibody directed specifically against an epitope close to the antigen binding site in B 1-8, the presence of aggregated B 1-8 produced by the plant was finally also detected.

By direct Western blotting from the crude extract of calli or induced plant material, only unsatisfactory results can be achieved. Only the detection of processed light chain was possible for different callus material. By using affinity chromatographic purification with NP Sepharose and then Western blotting, the presence of a functional aggregated antibody to be indirectly demonstrated from induced leaf material of a greenhouse A10 plant indirectly by detection of the light chain from isolated, biologically active hapten-bound plant B 1-8 was determined. Just beneath the band for the light chain a cross reaction unexpectedly occurred. Repetition of the experiment with A5 tissue with addition of CHAPS led to a cleaner but very weak band for the light chain from a larger charge of induced leaf material (50 mg total protein); the protein was purified both with Ls 136 (only light chain) and with NP Sepharose (functional aggregated antibodies). The signal of both analyses corresponds to processed light chain. In no case was such a band seen in the control with the wild type extract. By indirect detection of isolation of the functional antibody with hapten detection of the light chain contained in it, it could be demonstrated that plants are capable of synthesis of active monoclonal antibody whose individual chains are fused with a plant signal peptide.

The difficult reproducibility of biological material is particularly clear in these analyses, precisely when inductions must be carried out. The total found amounts of antibody protein lies in the lowermost range of the detection limits and therefore form only a very limited fraction of the total protein of the transformed plants.

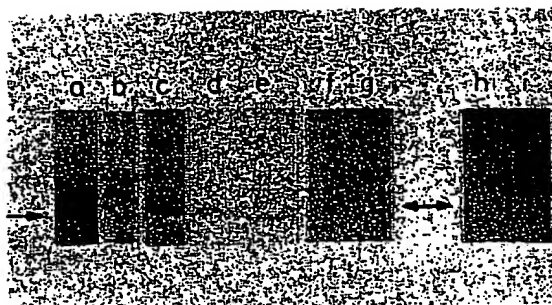


Figure III/23: Western blotting of transgenic A plants

a) Crude extract of A11 callus

b) Crude extract of W38 callus

c) Positive control: purified B 1-8

a-c) All tissue samples were extracted in denaturing probe buffer at 95°C, detection with biotinylated anti- λ antibodies

d) Detection of functional B 1-8 and induced greenhouse leaves (A10) by affinity chromatographic purification with NP Sepharose and denaturing Western blot from PBS extract, detection with anti- λ antibodies

e), like d) Negative control with W38 leaves

f) and g) Analysis of 10 mM CHAPS-PBS extract from induced greenhouse leaves (A5) by affinity chromatographic purification and denaturing Western blot, detection with anti- λ antibodies,

f) Detection of light chain with Ls 136 antibodies

g) Detection of functional B 1-8 with NP Sepharose

h) and i), like f) and g) Negative control with W 38 leaves.

3.15 Expression analysis of lysozyme plants by Western blotting

For analysis of lysozyme plants, the same techniques and principles were used as for the antibody plants and are described in 3.14.

For immunodetection of T4 lysozyme, a polyclonal anti-T4 lysozyme antibody was produced by injection of protein expressed in bacteria (furnished by R. Wetzel, Genentech). This was purified as described in 2.8. Despite purification via an affinity column, this antibody easily enters into cross reactions with the larger available amounts of nonspecific proteins and must therefore be appropriately pretreated.

By tissue printing (Cassab and Varner, 1967), both specific signals could be obtained, since despite all saturations and washings, coloring always occurs in the controls.

Direct detection of T4 lysozymes in crude extracts should be readily possible after saturation of the promoter activity in the calli and also in strongly active leaf material, but is hampered by the occurrence of many cross-reacting bands in the Western blot. Only very weak signals could be obtained with this method which convey specificity. From highly active calli, obtained by direct DNA transfer, very weak bands could be seen with the size of the T4 lysozyme applied as reference (produced in bacteria) (18.7 kDa), with a size of about 21 kDa and about 23 kDa. All intensities, however, lie at the outermost limit of the detection possibility and therefore indicate much lower presence of the sought protein than expected. As a result, processed and nonglycosylated T4 lysozyme must be present in the plant cells and the 21 kDa bands correspond to its precursor protein which still contains the signal peptide. In the extremely weak 23 kDa band, a glycosylated T4 lysozyme must be involved in which no assertion can be made concerning processing. Because of the low signal intensity, these bands can only be documented photographically with considerable difficulty.

The detection of T4 lysozyme by Western blotting after enrichment by the affinity chromatographic method did not lead to success either without detergent or with Triton or CHAPS, although high NPT activity had been found in similar tissue. Since the presence of the sought protein in the plant tissue could be demonstrated without doubt and the T4 lysozyme is still primarily localized in the intercellular space (see 3.17), the nonidentifiability of this protein in the Western blot after previous preparations of crude extract can only be attributed to insolubility, which cannot be eliminated by detergents, or to instability. On the other hand, the variability of expression from cell to cell determined in immunogold labeling can also lead to only a low overall frequency of T4 lysozymes.

Because of the difficult detection, no assertion can be made concerning the ratio of three found forms of T4 lysozyme (nonglycosylated-nonprocessed and processed, as well as probably a glycosylated form). The question is raised whether the glycosylated and nonglycosylated forms had different behavior relative to secretion.

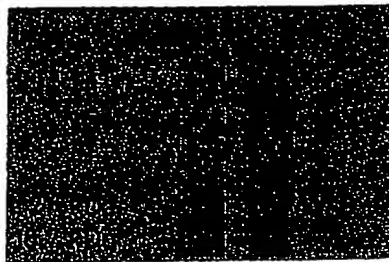


Figure III/24: Detection of T4 lysozyme in a callus obtained by naked DNA transfer from a CHAPS-SDS crude extract by detection with anti-T4 lysozyme antibodies.

- a) Transformed callus,
- b) SR1 wild type callus.
- c) Positive control: T4 lysozyme purified and produced in bacteria (furnished by R. Wetzel).

3.16 Localization of formed antibodies by immunogold labeling

All preparations and the immunogold labeling were conducted in cooperation with Dr. Sigrun Hippe (Aachen).

Modern low-temperature electron microscopy uses the specially developed Lowicryl synthetic resin from a few years ago for localization of membrane-bound and also dissolved compounds in biological tissue as in embedding agents (Kellenberger et al., 1980; Carlemalm et al., 1982). Lowicryl-K4M is a polar and Lowicryl HM20 a nonpolar embedding resin. The major advantage consists of the good infiltration capability of biological material and UV polymerization at low temperature. This mild preparation retains the ultrastructure of the embedded tissue excellently.

Comparison of Spurr's resin with Lowicryl HM 20 in preparation of suspensions of the phytopathogenic fungus *Ustilago avenae* during cryofixation in the propane jet demonstrates the improved structure retention (Hippe and Hermanns, 1986; Hippe, 1987). For the present problem, i.e., detection of foreign proteins present in relatively limited amounts in compact plant tissue (stem, root, leaf tissue), the preparation conditions had to be adjusted. The first attempt with the propane jet, which is well suited for cryofixation of plant cell monolayers (Hippe, 1985) did not lead to adequate retention of cellular structure. In addition, it had to be kept in mind that the sought protein is both bound in the cell and also secreted into the intercellular space, i.e., can be freely mobile. A conventional preparation at room temperature, like the PLT technique (progressive lowering temperature) is therefore ruled out, since both methods do not rule out

extraction of soluble cell and tissue components. Water-rich plant tissue can be frozen with the mentioned methods, as successfully used for the cell monolayer, but not optimally, since despite repeatedly modified attempts ice crystallization and the related partial destruction of tissue could not be prevented. A glass-like solidification (vitrification) of the tissue requires a shock freezing rate of 5.000-10.000°C/sec). By using high pressures, the required freezing rate can be reduced from the preparation. The high-pressure freezing method was developed by H. Moor (Zurich) in cooperation with the Balzers Co. (Balzers HPM 010) and this apparatus is the best one now available that meets the above requirements and optimally vitrifies the tissue samples up to 20 µm thick (Moor, 1987). All required probes for immunogold labeling were therefore prepared in the laboratory by H. Moor (Zurich) by means of the HPF method (high-pressure freezing) on the Balzers HPM 010. The structure retention found in the investigated preparations is excellent. For trouble-free localization of the sought proteins for detection of secretion, this is an absolute requirement.

After freezing substitution in acetone, low-temperature embedding in Lowicryl HM 20 resin and polymerization under UV light -35°C occur and thin sections are prepared at room temperature by means of a diamond blade and trapped on filmed and carbon-sputtered nickel grids. The immune reactions were adjusted to the experiences from Western blotting. All employed gold preparations came from the Janssen Co. The electromicroscopic analysis was then conducted (see 2.13).

For analysis of the A plants, it must be considered that the two genes for the light and heavy chains of the B 1-8 antibody are under the control of two different promoters. As mentioned in 3.6, the two promoters (pT_R and pNOS) can achieve roughly the same expression in induced leaf tissue and a double promoter can there possess roughly five to six times the strength of the nopaline synthase promoter under particularly favorable conditions. Greater differences, on the other hand, are possible in calli, because the double promoter experiences a significant activity increase there in contrast to the NOS promoter.

As a result, analysis of calli only made sense with detection of the light chain. For detection of intact antibodies, induced stems can be used because of the on average better structuring of the plant tissue relative to callus tissue.

The observed results are based on analysis of some sections of callus tissue with the LS 136 antibody (anti-λ) and induced stems with the Ac 38 antibody (anti B 1-8). The employed first and second antibodies, as well as streptavidin unexpectedly exhibited distinct nonspecific binding in the controls which is not the case in the polyclonal anti-T4 lysozyme antibody (see 3.17) and is therefore in contrast to the experiences from Western blotting (see 3.14 and 3.15). By using different blocking agents (gelatin, lysine, W 38 wild-type extract, goat normal serum),

however, cleaner controls could finally be obtained. This is an absolute requirement for clear interpretation of the labelings at the low signal frequency found in these analyses.

Ultrathin sections of induced stems of plant A5 were incubated with Ac 38 antibodies in gold-adsorbed anti-mouse IgG antibodies (G 10). The controls did not exhibit nonspecific labeling. Labeling in this preparation is associated with the endoplasmic reticulum and is surprisingly found in chloroplasts in labeling-free surroundings. Each chloroplast contains about 10 to 20 gold particles, often localized on thylakoid membranes. Sections of A5 callus tissue were incubated with biotinylated Ls 136 antibodies and then with streptavidin gold (G 10). Conspicuous labeling is found here in accumulations of 3-10 gold particles in the cytoplasm which such accumulations not being seen in the controls. In none of the analyses thus far has labeling been found in the vicinity of the cell walls or in the intercellular space or an indication of association of antibodies or the light chain with the Golgi apparatus or vesicles.

Further detailed analyses with both antibodies will provide information concerning the statistical significance of the observed signals and permit a more extensive localization.

By means of the available data it must be stated that synthesis and assembling of a monoclonal (B 1-8) antibody whose light and heavy chains are fused with a plant signal peptide occur in the plant cells on the rough ER and detection occurs with Ac 38 antibodies directed against the intact B 1-8. By incubation of the Ls 136 antibodies, the light chain could be detected both in isolated form and bound in aggregated B 1-8 antibodies. By comparison with the Ac 38-labeled preparation in which no free cytoplasmic labeling has thus far been found, it is assumed that at least partially isolated chains are involved which are localized freely in the cytoplasm. Aggregated B 1-8 antibodies are also found in the chloroplasts, since the relative percentage of gold particles bound there is very high relative to the total number. In comparison, this percentage is very low for T4 lysozyme (see 3.17). Together with the data from Western blotting, it is therefore clearly demonstrated that synthesis and assembling of a monoclonal antibody is possible in plants, at least during fusion of the individual chains with a plant signal peptide.

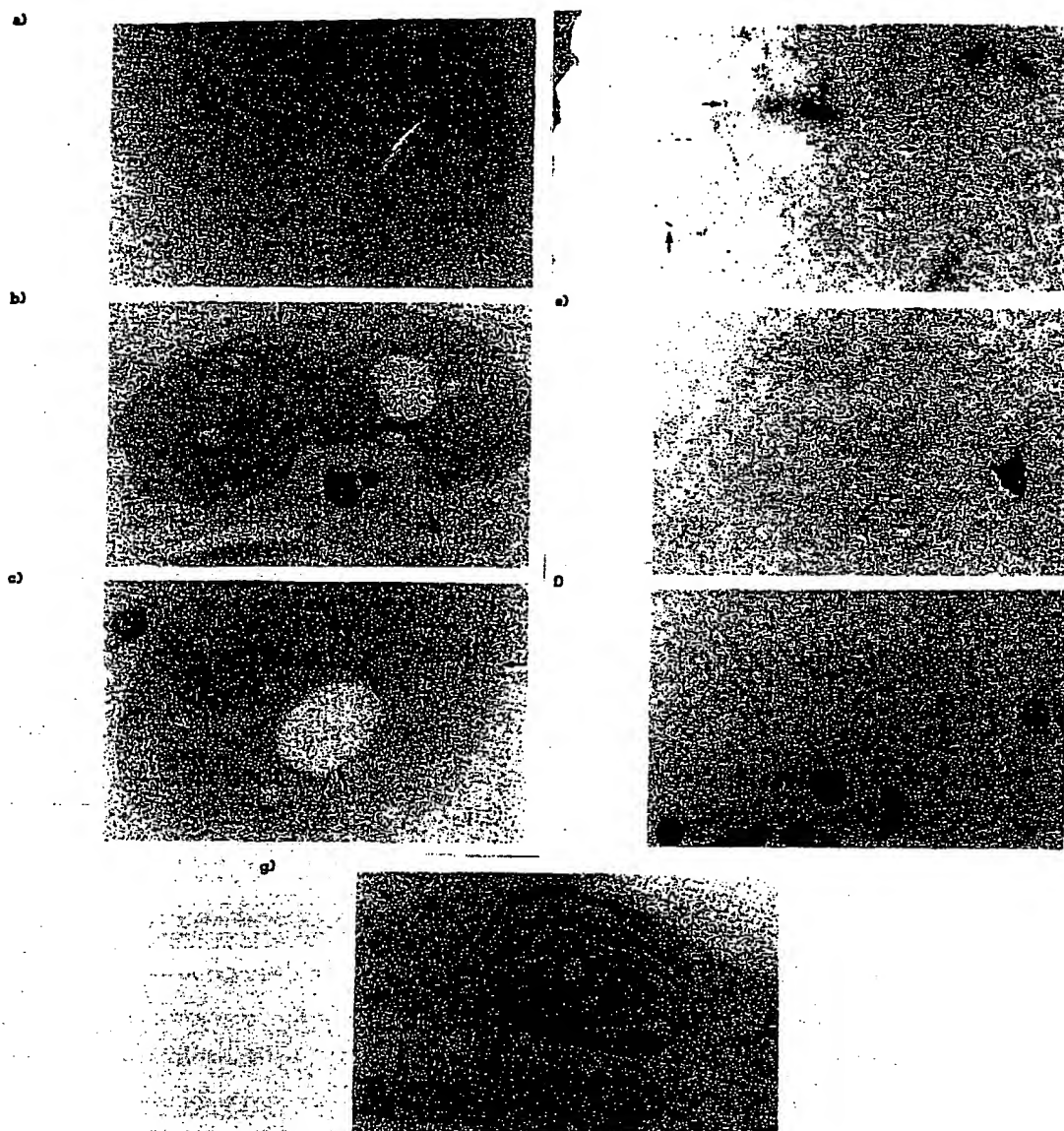


Figure III/25: Localization of B-1-8 antibodies and their light chain by immunogold labeling. A5 and W 38 wild type stem tissue (induced) and callus, prepared according to 2.13, labeling either the Ac 38 and gold-adsorbed anti-mouse IgG antibodies (G 10) (a-c, f-g) or with biotinylated Ls 136 and streptavidin gold (G 10) (d-e).

a) A5 callus: endoplasmic reticulum with labeling (1:35,000/Ac 38)

b) + c) A5 lowermost stem piece induced: typical labeled chloroplasts (1:18,000 to 1:26,000/Ac 38)

d) + e): A5 callus: accumulated cytoplasmic labeling (1:38,500 to 1:48,000/Ls 136)

f)-g): W 28 wild type control

f) Cytoplasm with endoplasmic reticulum (1:17,000/Ac 38)

g) Chloroplast (1:30,000/Ac 38)

3.17 Localization of formed T4 lysozyme by immunogold labeling

With the method of low temperature-high pressure preparation described in 3.16, calli and leaves, fresh stems and induced ones, as well as roots, from the plant L2 were frozen. The most promising samples after the NPT test of adjacent tissue pieces were both the induced stems and leaves and the fresh parts and calli: the highest activity was exhibited by the induced stems and these samples were therefore chosen for first examination. In *Acetabularia*, the anti-T4 lysozyme antibodies used here were already successfully used to detect T4 lysozyme produced there in indirect immunofluorescence. Because of this analysis, it was known that production of a protein that is recognized specifically by this antibody and therefore must be identical to T4 lysozyme is possible in plant organisms. Since localization of foreign proteins in transgenic plants has still not been described in the literature, the degree of expression in relation to the expected signal was an uncertainty factor. Moreover, different antibody systems produced differently strong signals with the same presence of antigen. The highest expressing plant parts were therefore initially used for investigation.

In samples of induced stems, alternating strong signals were found for T4 lysozymes. By corresponding blocking of nonspecific binding types of this antibody with wild type W 38-PBS extract, cross reaction can be ruled out. In the controls only in normal, extremely limited and statistically distributed frequency of nonspecifically bound gold particles occurred. The promoter intensity obviously varies from cell to cell, since both strongly expressing and weakly expressing cells are found in one preparation in the series section. This can be attributed to the worked out (see 3.6) hormone-dependent expression of the double promoter, if one assumes that the hormone ratios can vary from cell to cell. However, other factors possibly also play a role here, including damage of the cells by produced T4 lysozyme.

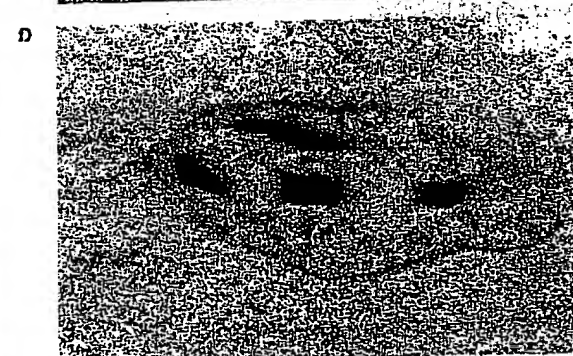
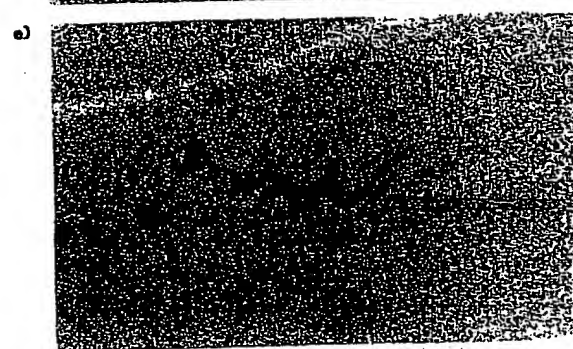
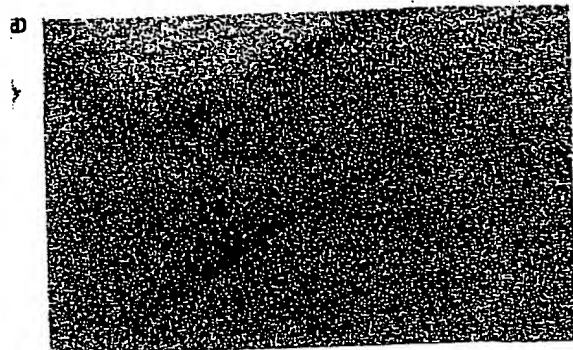
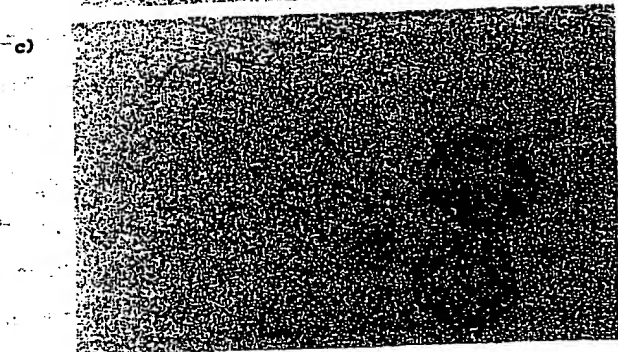
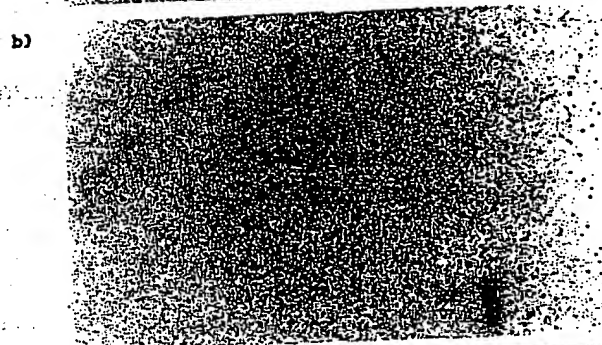
Immunogold labeling (see Figure III/26) is found in strongly concentrated form in the intercellular spaces of the stem cells, while the gold particles are associated with filament-like, network-like structures. They start sometimes from the cell wall and enter the interior of the intercellular space and sometimes branch off and, on the other hand, are present in isolated fashion in clusters in the interior. The structures are very likely associates of either pure T4 lysozyme or are in conjunction with other proteins and are not found in the controls. Significant numbers of gold points are found in isolated form in the cytoplasm. They are situated on filamentary structures, mostly on the periphery of the cell. Characteristic accumulations of gold labelings occur in and in the vicinity of plasmodesmata and in this way the synthesized T4 lysozyme is distributed within the cells. Labeled structures are observed differently in the cytoplasm, which must be viewed as crystallized lysozyme. Gold labeling was strikingly detected in series section in proplastids of the phloem on strongly electron-dense, filamentary or crystalline structures. The crystalline structures in the cytoplasm and in the proplastids are

similar to each other. Weak labeling can also be observed in the chloroplasts. A weak labeling is also found in isolated fashion in the cell walls. The gold points are homogeneously distributed in the phloem cells in the vicinity of the cell wall and transported T4 lysozyme occurs here through the entire plant. In particular, cells in the region of labeled lysozyme-rich intercellular areas exhibit intracellular labeling.

The production of T4 lysozyme in plant cells was therefore clearly confirmed. The protein is obviously not produced cytoplasmically, but in secretory fashion and secreted by the endoplasmic reticulum. The largest part of the produced T4 lysozyme is localized intercellularly, but part is obviously stored by the plants in the phloem proplastids and also cytoplasmically.

Gubler et al. (1987) revealed, with immunofluorescence and also immunogold labeling, the localization of α -amylase in aleurone tissue in the resting state and after induction with gibberellic acid. Previously the same labeled regions were described there: fluorescence in the cytoplasm at discrete sites, accumulated in some regions, but of the same type in all cells, after induction also in the vicinity of the internal cell wall. The cytoplasmic fluorescence occurs in special strand-like structures, generally in the periphery of the cell; the signal in digested cell wall regions are associated with the plasmodesmata. By immunogold labeling, strong hydrolysis of the cell wall and labeling around groups of plasmodesmata could be refined and a larger number of gold points could be detected in the vicinity of the plasmodesmata in undigested wall pieces. Localization in proplastids or chloroplasts has thus far not been reported. It can be stated as a result of immunogold labeling on stems of the L2 plant that the specificity of localization of α -amylase from barley in aleurone tissue is also retained for the signal peptide used with T4 lysozyme in the stem tissue of tobacco plants. The T4 lysozyme is also found in the intercellular space and has therefore passed through the cell wall and has been efficiently secreted. The size of a globular protein, which can still pass through pores of the cell wall is estimated at about 3.8-4.0 nm diameter with a molecular weight of about 17 kDa (Gubler et al., 1987), and T4 lysozyme has a molecular weight of 18.7 kDa.

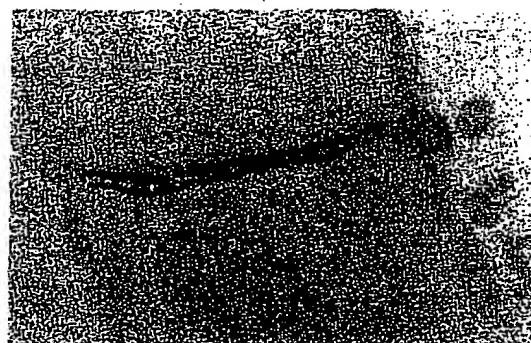
The objective of this project, to permit secretion of the protein in plants by fusion of T4 lysozyme of a prokaryotic protein to the signal peptide of a plant protein, was therefore achieved. Passage through the cell wall is obviously made possible by the still suitable size of the protein, which in the case of α -amylase is no longer possible and occurs only by gibberellic acid-induced hydrolysis of the cell wall. Characteristic details of secretion of α -amylase, which were only recently described, are also maintained in these plants produced by genetic engineering, which permits interesting conclusions concerning the method of functioning of signal peptides, at least in plants. The localization of the foreign protein in obviously crystallized form in proplastids of the flowing region is an important interesting secondary phenomenon.



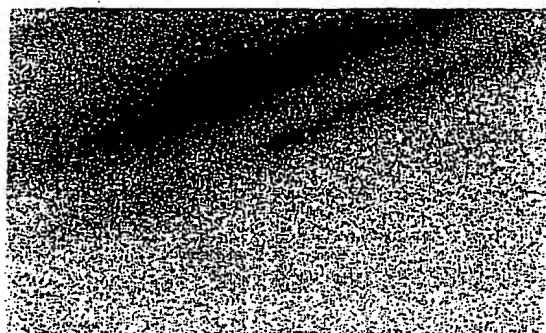
g)



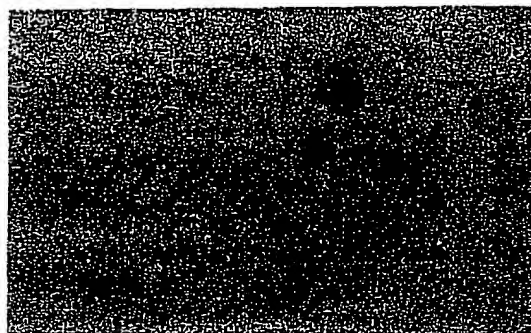
k)



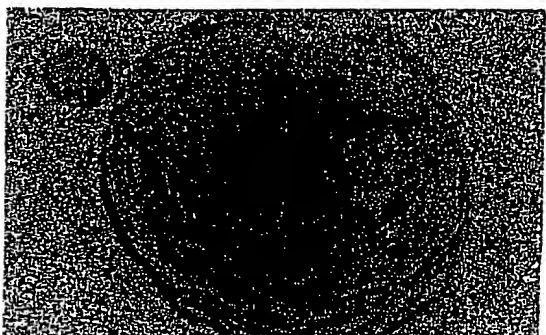
h)



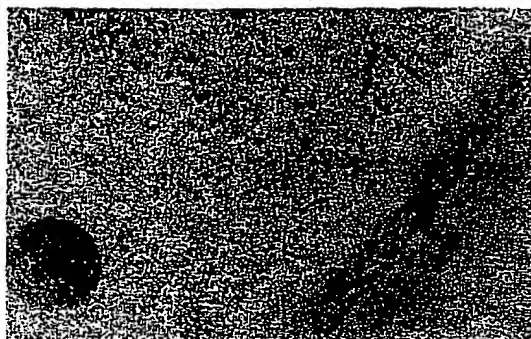
D



D



m)



n)



Figure III/26: Localization of T4 lysozyme by immunogold labeling.

L 2 and W 38 wild type stem tissue prepared according to 2.13, labeling with anti-T4 lysozyme antibodies and protein A gold (G 10).

- a)-i): L2 lowermost stem piece
 - a) Intercellular space with strong gold labeling (1:14,000)
 - b) Localization of T4 lysozyme on filamentary network-like structures in the intercellular space (1:64,000)
 - c) Labeling in cytoplasm (1:41,000)
 - d) Labeled needle-like crystals in cytoplasm (1:62,500)
 - e) + f) Proplastids with labeled crystalline inclusions (1:21,000 to 1:35,000)
 - g) Chloroplasts with labeling (1:16,000)
 - h) + i) Concentrated accumulation of gold particles in the cell wall region of phloem cells (1:43,000 to 1:35,000)
 - k) Detection of T4 lysozymes in plasmodesmata (1:34,000)
- l)-n): W 38 wild type control
 - l) Section from a cell (1:24,000)
 - m) In cytoplasm section (1:41,000)
 - n) Intercellular space (1:18,000).

IV. Discussion

The task posed in the present study of expression, assembling (in antibody) and secretion of single and multiple-chain proteins not endogenous to plants in *Nicotiana tabacum* could not be solved without developing modern highly sensitive detection methods and carrying out activity analysis of the employed promoter.

All constructions were based primarily on the use of T_R -1',2' double promoter (Velten et al., 1984; Velten and Schell, 1986), which at the time of construction was considered as a constitutive, strongly expressing promoter and had been constructed from the selection-expression vectors (Velten and Schell, 1986). The neomycin phosphotransferase II (NPT) gene is available in the 2' position as selectable marker and reporter gene. The relative intensity of expression of the gene being tested cloned in the 1' position can be checked with this. Thus far it was only known that both promoter directions are expressed equally strongly in calli (Velten et al., 1984; Velten and Schell, 1986). M. Prols and R. Topfer (personal communication) in the mean time also tested transient expression in carrot protoplasts and here also found equivalent activity of both directions in comparable intensity of this promoter with the CaMV 35S promoter that was then recognized as the strongest constitutive promoter, i.e., about 0.1% of the specific protein in total protein. The same applies for expression of these two promoters in calli.

The T4 lysozyme gene and the gene for the light chain of the monoclonal antibodies were cloned in the 1' position of the vector pAP 2034 (Velten and Schell, 1986). For the heavy chain of the monoclonal antibody, the nopaline synthase promoter from T-DNA of *Agrobacterium tumefaciens* was chosen as the second available promoter. The genes for both chains were cloned into a plasmid.

After a few RNA and protein analyses were conducted without success, there were indications that the T_R double promoter is not expressed constitutively, but in differentiated fashion in intact plants and also through induction by plant hormones (A. Szalay, C. Koncz, personal communication). An activity analysis of this promoter was then undertaken in order to clarify under which conditions expression is to be expected at all and which activity overall can be achieved in the different tissues and under which culture conditions. The starting point for the induction analysis was the assumption that the promoter is inducible by auxins and/or cytokinins (A. Szalay, C. Koncz, personal communication). As described under 3.6, a series of different hormone concentrations and ratios were tested under different conditions and with different tissue parts for induction, as well as fine-structure for all possible parts of the sterile culture plant with the NPT test to their relative activity. Comparison for quantitative estimation of the expression level was plotted in each case with the extract of the 35S NPT callus. The following picture emerges as a summary of all these experiments: the double promoter is expressed in the intact plants increasingly from the top down in an activity gradient. Tissue-specific expression cannot be spoken of, since higher signals are not only measured from the roots, but also from the lower part of the stem and sometimes also from the lowermost leaf. The expression pattern in the different regenerated transformants, which all start from the same vector pAP 2034, is very different. A large part exhibits the highest activity in the transition region from the stem to the

roots, another part, however, exhibits higher activity only in the roots. Only sometimes do we also detect good activity in the lowermost leaf (especially on the front edge), but only slightly higher expression prevails here than in the other leaves. The differentiated activity distribution obviously also depends on the development state of the investigated specimen of a plant. The increase in activity in the double promoter in the transformed plants from the top down qualitatively agrees with the previously known data from A. Szalay.

The studies on inducibility of the promoter led to a positive result. It can be achieved both by wounding alone and by addition of plant hormones. 2,4-D and kinetin 1:1 in a 10^{-5} molar concentration proved to be the optimal hormone combination. These conditions were tested by incubation on the base of cutoff leaves in liquid M + S medium with and without additives. In this case, efficient induction is obtained only with hormone addition. The promoter behaves differently when the leaves are cultured on solid medium with and without additives cut into small pieces. No difference can then be found between the two conditions and both times the same induction is achieved. If cut stem segments are used for induction, the same effect without differences is obtained. It must be concluded from this that the induction factors depend on the ratio of wound surface to total tissue. In whole plants, this ratio is extremely low, in cut stem segments it is highest and the inducibility is increased accordingly with an increase in this ratio from whole leaves via leaf strips to leaf pieces. Generally all tissues, whether from the uppermost or lowermost leaf, are capable of achieving the same level of expression.

Extending these results, the very high expression in calli (which are cultured on exogenous hormones) and the good transient expression in protoplasts fit with the previously described data. Calli are the natural environment in which after infection of wounded plants the opine genes of the transfer T-DNA from *Agrobacterium tumefaciens* are activated (Gelvin et al., 1981). The hormone ratios and thus growth of this crown gall is coded and controlled by the genes 1,2 also coded by T-DNA (for auxin synthesis) (Schroder et al., 1984; Kemper et al., 1985) and gene 4 (for cytokinin synthesis) (Akiyoshi et al., 1984). The morphology of such tumors is determined by the amount and ratios of the produced plant hormones (Amasino and Miller, 1982; Akiyoshi et al., 1983). This gives a further indication of the effect of auxin and cytokinin on expression of the double promoter. The observed differently high expression in the calli of different morphology fits in conclusively: from unorganized to higher organized states, expression diminishes, while the hormone ratios in natural tumors of the same texture (Amasino and Miller, 1982; Akiyoshi et al., 1983) deviate from the experimentally determined optimal values. Unorganized calli possess comparatively lower percentage of auxin and roughly the same amount of cytokinin as well organized calli. These measurements, as well as the fact that the highest expression in the plant lies not at the tip, but in the transition region from the stem and roots, confirm the experimental finding that cytokinin is also necessary for induction.

An interesting study in this respect is the recently published study concerning the stability of cytokinin in the presence of auxin (Palni et al., 1988). The authors found an inverse correlation between the concentration of auxin and cytokinin in plant tissue. Degradation of cytokinin occurs at least partially through the enzyme cytokinin oxidase whose activity is obviously influenced by the concentration of auxin. The percentage of cytokinin taken up in the plant tissue of the total supplied cytokinin is about 30%. The ratios described in this study offer a basis for further detailed mechanistic studies on regulation of the T_R double promoter. The effective hormone concentrations and ratios in the plants actually appear to be regulated in far more complex fashion than would correspond to the exogeneously applied conditions. A specific analysis of the background of these data should make it possible to establish the hormone conditions that are ultimately optimal for activity of a double promoter.

In addition to *in vitro* inductions in sterile culture, an induction on the intact plants could also be achieved on greenhouse plants by wounding and hormone treatment. It could therefore be assumed that the same mechanisms are activated and the double promoter is induced *in vivo* during wounding from, say, pests.

The inducing factors naturally could not be clearly identified based on these experiments, since the objective was only to work out optimal expression conditions. The effect of wounding and of exogenous hormone addition was clearly demonstrated, but it is not clear whether the substance acetosyringone synthesized by a wounded plant cell exerts an effect. The fact that wounding alone is sufficient for activation of this promoter was demonstrated by numerous experiments with stem segments and by cutting of leaves into small pieces. In these experiments, no additional increase could be achieved by additional hormones. With a very small wound surface, however, induction only occurs by addition of hormones. Endogenous factors alone are obviously capable of driving the activity of the promoter to the highest possible stage and these conditions apparently cannot be simulated from the outside. However, wounding and hormone effects are also involved here.

Wounding and hormones are each sufficient alone for the exogeneously highest possible attainable induction of the double promoter, but both together do not yield an additive effect. It therefore cannot be stated which of the two is the primary or secondary effect. It is known that during wounding hormone biosynthesis is activated, for which reason it might be that wounding is merely a triggering factor and the forming hormone concentrations and ratios are the regulating factor. This hypothesis is supported by the finding that differential expression from the top down is also present in intact healthy plants. Auxin biosynthesis occurs in the tip of the plant (apical dominance) and that of cytokinins in the roots. Both hormones are transported oppositely through the plant so that a hormone gradient is produced. The apical dominance of the plant is eliminated by cutting off the tip, the hormone conditions are altered and the activity of

the T_R double promoter rises (A. Szalay, personal communication). All these observations suggest direct regulation by hormones, however, in which the effect of other (especially wound-dependent factors, for example, acetosyringone) are not ruled out, especially for the highest possible expression.

The results from immunogold labeling show that the activity of the double promoter obviously can also deviate very sharply in neighboring cells. Cells with very high expression and cells with weaker expression were found next to each other. This phenomenon might be due to different states of development of the individual cells and therefore opens up additional approaches for analysis of promoter regulation. In the case of lysozyme plant L2, this effect can also be attributed to effects of the produced lysozyme of different intensity.

The single lysozyme plant L2 obtained from a leaf disk test shows an abnormally high base level of expression in intact tissue in all NPT tests and, unlike all the other analyzed transformants, is no longer inducible. At the wounding sites on the lower stem, hard, white (well organized) calli repeatedly form, which were sometimes subjected to an NPT test. As expected, they exhibited about 1/5 the activity of unorganized soft calli that correspond to the activity of a 35S NPT callus. A possible explanation would be damage to the cell wall by T4 lysozyme synthesized by the plants, which could cause a persistent state of wounding (see below). This demonstrates that the induction conditions worked out *in vitro* also occur *in vivo* without external additives and lead to the same results.

All previously discussed results refer to analysis of the 2' position of the double promoter. In order to also test the asserted and thus far detected equally high expression of the 1' position only in calli, only transformed plants could be resorted to that contain the 1' NPT gene in an environment other than the previously investigated 2' NPT gene (furnished by 'G. Coupland, Cologne). No general differences in expression and inducibility were found for these plants. However, a systematic study with a construction that makes both directions in a plant analyzable and correlatable is necessary and being set up (R. Masterson, personal communication).

Transfer of the tests just described to the NOS promoter (Depicker et al., 1982; Bevan et al., 1983), despite all previous publications that describe it as clearly constitutive (see Sanders et al., 1987), was successful. Such a relation does not appear illogical, since both promoters control the genes for synthesis of opines and both are components of the T-DNA of the Ti plasmids from agrobacteria. The magnitudes of the maximum inducibility in the NPT tests differ clearly (pT_R up to 170 \times and 1000 \times , $pNOS$ up to 20 \times), but the factors for exogenous induction as well as the base states of expression of both promoters are of the same order.

Harpster et al. (1988) just published a comparison of CaMV 35S 1', 2' and NOS promoter on the RNA level by primer extension analysis. However, this study was only conducted in calli. Differently strong signals for the 1' and 2' position of the double promoter were found in tobacco

calli. The 35S promoter as a result is about two to three times as strong as the 2' position, but about 10 to 15 times stronger than the 1' position and for the NOS promoter only slightly lower values are found. The promoter strengths are therefore arranged in two groups, in which the 2' position is somewhat weaker than the 35S promoter and the 1' position as well as the NOS promoter are about equally weak (10% of the 2' activity). The found ratios for 35S and NOS promoter do correlate with those of Sanders et al. (1987), but the values for the double promoter are not consistent with the already known values and the ones measured in this study from the enzyme test. The measurement might also be influenced by the regulated expression of the double promoter, which is not considered, since no culture conditions are stated and no conclusions can be drawn concerning this. This again demonstrates the importance of activity analysis of both promoter directions in one transgenic plant and callus and protoplasts cultivated from it.

The molecular regulation mechanisms require a detailed analysis, which is being conducted for the T_R promoter by R. Masterson (personal communication). The differences between pNOS and p T_R suggest that the promoters could differ by the number of activating elements. Whereas the effectiveness of a conserved element that both possess could be imagined for the common regulation mechanism, the presence of a second activating element that must be regulated by additional effects would be conceivable for the also strong activation in calli of protoplasts. For an analogous 2' promoter of T_R -DNA from pTiB₈₀₆ (another Ti plasmid), DiRita and Gelvin (1987) found the presence of two motifs from the SV 40 enhancer (Zenke et al., 1986), which contain precisely the same sequences in the 1',2' double promoter analyzed here. They also identified a fragment of the promoter which has a regulating effect and can reactivate an octopine synthase minimal promoter. Even in the NOS promoter, a more precise structural investigation has since been conducted. It turned out that a 19-bp deletion led to complete disengagement of the promoter and duplication of this sequence, on the other hand, increased the promoter strength significantly (Ebert et al., 1987). Both promoters therefore contain activating elements.

The worked out properties of the T_R double promoter make it appear as an appropriate candidate for construction of resistance genes which are optimally engaged only as needed in order to protect healthy transformed plants from loading by resistance proteins. Thus far no other promoter has been described that has similar or more favorable properties. Mostly local strong activation of transcription occurs from wounding, which should be sufficient to make the necessary amounts of resistance protein available. The question whether activation occurs quickly enough has still not been clearly answered. In kinetic studies of in vitro induction of the double promoter, a rise in activity was found only after 1 to 2 days and the maximum expression after about 5 to 6 days. The data from greenhouse plants (stronger rise already after 1 day) and of

transient expression in protoplasts, however, suggest that more rapid induction occurs in vivo. By multiplication of activating elements, the inducibility should be further increased, as was recently demonstrated for the 35S promoter (Kay et al., 1987; Ow et al., 1987).

Optimal adjustment of the agrobacteria to their host is the specifically activated expression of bacterial genes that lead to synthesis of compounds in the wounded plant tissue that can be used as a raw material source by genes coded also on the Ti plasmid.

In addition to expression in plants, the double promoter was also tested for its expression capability in bacteria. Gelvin et al. (1981, 1985) found for the aforementioned analogous 2' promoter a weak activity both in agrobacteria and in *E. coli*. It is sufficient to culture these transformed bacteria under kanamycin selection. The double promoter analyzed here is also active in both bacteria, but the strength is not sufficient to perform selection by kanamycin. The principal difference of the two promoters is the size of the employed promoter. Whereas Gelvin et al. cloned the marker gene instead of the originally present mannopine synthase gene in the Ti plasmid, here we worked with an isolated 470 bp long fragment, to which foreign sequences are linked in both directions in vitro. Important signals in the complementary strand of the gene read in the other direction might be present for expression in bacteria.

The expression of a T4 lysozyme gene under the control of the 1' position in agrobacteria must therefore be carefully considered. It was possible without problems to introduce the plasmid pSR 2-4 by homologous recombination intact in agrobacteria. As a result, expression in a normal bacterial nutrient medium cannot be so high that these are lysed. Another explanation would be the inactivity of T4 lysozyme relative to agrobacteria. The activity of T4 lysozyme against *Micrococcus lysodeicticus* was demonstrated in the course of this work (see 2.12.11). The first explanation therefore appears more likely. In addition, an attack of cytoplasm on the cell membrane corresponds to the natural mechanism of infection of *E. coli* by the phage T4. Additional hints are obtained from transformation of *Nicotiana tabacum* with plasmid pSR 2-4. As already mentioned, only a single transformant could be regenerated from a number of leaf disk tests which incorporated the lysozyme gene. In this method, the DNA transfer is produced by agrobacteria, but, on the other hand, naked DNA transfer is carried out by means of protoplasts and a routine number of transformed calli is obtained (R. Hain, personal communication). No difference is found relative to transformation of other genes and consequently the problem in the first method must lie in stronger expression of the double promoter at the time of coculturing of the agrobacteria with the leaf pieces, which must result in increased lysozyme production and thus corresponding damage to the cell wall. As a result of this, DNA transfer is then inhibited. A negative effect of T4 lysozyme on the plant cell wall cannot be ruled out either. The plant L2 always has a very high NPT expression and is not inducible, which means the maximum expression is obviously already reached. This could be

attributed by the fact that the formed lysozyme exerts a degrading effect on the cell wall, which causes the liberation of signal molecules (similar to acetosyringone) and as a chain reaction and stimulation of the double promoter. Identification of the regulating substances of the double promoter will be helpful but the explanation of such a complex event as wounding of a plant cell with all the occurring reactions remains difficult.

A conclusive hypothesis can be advanced with the observations just made and demonstrated by our own experimental observations and other published data. Expression of lysozyme does occur in agrobacteria but only in nontoxic amounts. Matthyse (1986) was able to demonstrate that the presence of auxin is necessary in the bacterial cells for successful attachment of agrobacteria in plant cells. The auxin can either be produced by the bacterium itself or secreted by the wounded plant cell and taken up by the bacterium. Since the T-DNA genes for auxin biosynthesis (genes 1 and 2) are deleted in the Ti plasmid used here, only the auxin liberated by the wounded plant cell can consequently be considered a determining factor. As was demonstrated in this study, auxin alone also has a stimulating effect to a certain degree on expression of the T_R double promoter in plant cells. It can be assumed that a similar regulation is also possible on bacteria, provided that all necessary interacting factors are present there. As a result of increased expression of the double promoter, the content of T4 lysozyme in the bacteria will increase, which results in damage or even lysis. An alternative is possible expression of chromosomal bacteria auxin genes induced by plant wound secretions. The binding of *A. tumefaciens* to plant cells obviously involves bacterial surface proteins and lipopolysaccharides whose production is apparently also auxin-related (Matthyse, 1986). Transfer of agrobacteria DNA to plant cells obviously requires a complex machinery of bacterial and plant factors, some of which can be destroyed by T4 lysozyme. The single transformed plant obtained by the leaf disk test for plasmid pSR 2-4 must therefore be considered a random exception. These conclusions are supported by the experimental result that during a direct DNA transfer via protoplasts, no changes are observed relative to routine transformations. Unusually high expression of the double promoter was found in all NPT tests of plant L2, but on the other hand, induction of expression could not be achieved in any case in this plant. This suggests that this plant is also in intact form in a "wounded" state that offers favorable expression conditions for the T_R promoter. By secretion of T4 lysozyme, this enzyme reaches the cell wall and the intercellular space and it cannot be ruled out that similar bonds are present in the murein sacculus of bacteria that can be opened by lysozyme (lysozyme cleaves the glycoside bond between N-acetylmuraminic acid and N-acetylglucosamine). Attack can therefore occur on the cell wall from both sides. A state of wounding is produced by this that does leave the cell itself intact, but nevertheless activates wounding reactions and consequently also leads to the production of plant

hormones. This means that high expression of T4 lysozyme in plant cells would be ensured in a circular process by "artificial" autoregulation.

The expression of T4 lysozyme was detected in *Acetabularia* by immunofluorescence in cooperation with G. Neuhaus (Ladenburg). A clear assertion concerning secretion of this protein could not be obtained. No data were known concerning possible damage to *Acetabularia* cells.

Clear detection of the presence of T4 lysozyme in the transgenic plant L2 could be obtained by immunogold labeling. The percentage of this protein in the total protein appears to fluctuate sharply, which is probably attributable to specific properties of the T_R double promoter (see above). However, cell-specific expression can also reflect effects of T4 lysozyme produced particularly strongly in these cells and an autocatalytic mechanism (see above). The observed localization of T4 lysozyme confirmed the secretion of a monomeric small foreign protein under the control of a plant signal peptide expected of the gene fusion constructed in it this work (pSR 2-4).

The signal peptide of α -amylase from barley was selected at the beginning of this work because at that time it was the best investigated secretion system in plants (Jacobsen and Knox, 1973; Jones and Chen, 1976; Locy and Kende, 1978; Jones and Jacobsen, 1982). Processing and secretion of intact wheat α -amylase in yeast had already also been successful (Rothstein et al., 1984). A cDNA clone (clone E) was previously isolated and characterized (Rogers and Milliman, 1983). Since then, additional very important data concerning secretion and location of α -amylase from barley in aleurone tissue had become known (Fernandez and Staehelin, 1985; Gubler et al., 1986, 1987).

Fusion of a foreign protein to the plant signal peptide obviously has no effect on its functional capability and specificity. The secretion mechanism, at least by T4 lysozyme, is not hampered or adversely affected. This can naturally be attributed to the favorable properties of this protein. Production of this protein in tobacco cells presumably does not occur on free ribosomes, since it was detected by gold labeling that a secretion pathway is involved. Gold particles are found in some regions of the cytoplasm on filamentary, strand-like and network-like structures. According to the current interpretation, a conformation must be maintained by the signal peptide that permits insertion of the newly synthesized protein into the ER membrane. If a protein is folded before insertion to its final spatial structure, no insertion is possible any longer in most cases (Zimmerman and Meyer, 1986) and only a few exceptions are known (Muller and Zimmerman, 1987; Walter, 1987; Wiedmann et al., 1987b). No membrane-anchoring sequences (Mize et al., 1986; Munro and Pelham, 1987) can be present in T4 lysozyme either (stop transfer sequences). No assertion can be made by means of electron microscopic data concerning processing of the precursor protein during entry into the ER membrane. The natural conserved positions are no longer present in this fusion, since no parts of the mature α -amylase protein

were transferred. Instead, a direct linking of the last amino acid of the signal peptide to the first amino acid of T4 lysozyme was constructed. However, secretion is also possible if no processing functions occur (Kaiser et al., 1987); the remaining additional amino-terminal polypeptide part can then have negative effects on the functional capability of the synthesized protein. The available data from Western blotting, however, show that processing of the chimeric protein obviously occurs.

T4 lysozyme has a potential consensus sequence for N-bound glycosylation (Asn-X-Ser/Thr) in natural prokaryotic systems; no glycosylation occurs so that this sequence has no significance there. It is localized between amino acids 140 and 142, which are situated in a position that is well accessible from the outside (cf. Remington and Mathews, 1978; Mathews et al., 1981, 1983). No steric hindrance for glycosylation in the ER and Golgi apparatus should therefore be expected. A change of the protein can occur by possible glycosylation in the plants, from which, however, no effects on the functionality of the active center are expected, since both regions are adjacent but are not functionally linked (see Mathews et al., 1983). On the other hand, examples are known in which glycosylation is essential for secretion (Olden et al., 1982; Ferro-Novick, 1985; Guan et al., 1985). Glycosylation of the produced T4 lysozyme could not be analyzed because of the limited detectable amounts.

The fact that the labelings in the experiment with the transgenic plants occur at the same sites as found by Gubler et al. (1987) for the aleurone tissue of barley with α -amylase itself, the association with filamentary structures in the cytoplasm in concentrated form in some regions of the cell and the localization on and in the plasmodesmata should be emphasized in particular. These characteristic details of secretion of α -amylase in completely different tissue of another plant are retained during genetic engineering construction with a foreign protein in tobacco. Transport of α -amylase into the vicinity of plasmodesmata obviously occurs because a highly specialized cell wall region exists in aleurone cells that differs distinctly in composition from the rest of the cell wall. This internal resistant wall layer preserves its integrity even in the presence of strong cell wall-hydrolyzing enzymes. Cell wall hydrolysis occurs most often in the vicinity of plasmodesmata, for which reason the most favorable conditions for secretion of proteins through the cell wall are present here. These structures are not present in normal tobacco cells, for which reason such specialized accumulation of the secreted protein would not be necessary there and as a result should not be coded by the cell either. Nevertheless, the same properties are observed, which can only be attributed to the fact that specific information for localization of α -amylase must be stored within the signal peptide, since these are the only consistent sequences between natural α -amylase in the signal peptide-T4 lysozyme fusion. Thus far, no clarity has yet been obtained whether such information is stored in a signal peptide or whether specific localizations are coded within other sequences in mature cells and the single task of the signal peptide is

merely insertion of the nascent protein in the ER. The present results clearly suggest that, at least in the case of α -amylase from barley, such information must be stored in the signal peptide.

After the T4 lysozyme has passed through the entire secretion path of α -amylase and has reached the cell wall, mostly in the vicinity of plasmodesmata, the differences between the two proteins now come to bear. α -Amylase is an enzyme with a molecular weight of about 45 kDa, whereas T4 lysozyme at 18.7 kDa is much smaller and has an ellipsoidal structure (cf. Grutter et al., 1983). The maximum diameter of a globular protein that can still pass through pores in the cell wall is estimated at about 3.8-4.0 nm, which corresponds to about 17 kDa (cf. Gubler et al., 1987). Since T4 lysozyme, however, has an ellipsoidal structure, it must still fall in the permitted range and passing through the pores of an intact cell should be possible. Gold labeling in the intercellular spaces of the transgenic plants is actually also found, i.e., secretion of T4 lysozyme is also possible through the cell wall up to the intercellular space by analogy to the pathogenesis related (PR) proteins (Parent and Asselin, 1984; Carr et al., 1987).

By fusion of a plant signal peptide with a nonsecreted small prokaryotic protein, it is also possible to achieve its extracellular localization. Additional studies on detailed analysis of the secretion pathway can be conducted on the present preparations (localization in ER, in the Golgi apparatus and in vesicles, etc.).

The found inclusions of T4 lysozymes in obviously crystalline form in proplastids raise the question whether transformed plant cells are eliminated by deposition of foreign protein (sometimes also observed cytoplasmically). Localization in chloroplasts, which repeatedly occurs, can be caused either by transport into fully developed chloroplasts or by the presence of T4 lysozyme in proplastids, which can later develop sometimes to chloroplasts. The first possibility appears to be quite unlikely because of the significant difference between signal peptides for secretion and transit peptides for chloroplast transport. Such a question can be answered by cloning T4 lysozyme gene in an SP 6 plasmid and in vitro translation with a subsequent test for incorporation in the isolated chloroplasts. The assumption that the gold labeling in the chloroplasts can be attributed to development of lysozyme-containing proplastids in the chloroplasts is more likely, although storage proplastids are commonly present in phloem tissue that can no longer be differentiated. It remains to consider that such data were found with even greater relevance for the antibody construct. The question how and why lysozyme reaches the proplastids remains open. This protein, which must pass through a membrane in bacteria for lytic dissolution of the murein sacculus, might have the capability of passing through other membranes. A relation between membranes of proplastids and bacteria is speculative, but cannot be ruled out on evolutionary grounds.

The different behavior of lysozyme with respect to secretion and cytoplasmic localization might be attributable to the presence of glycosylated and nonglycosylated forms.

Since T4 lysozyme is to be investigated as an antibacterial resistance protein in plants, extracellular localization in the intercellular spaces is the optimal solution. Resistance investigations with the plants regenerated by naked DNA transfer for T4 lysozyme are being prepared in cooperation with the Institute for Plant Diseases (Dr. Kuck) of the Monheim Plant Protection Center of Bayer AG.

The difficulties in detecting T4 lysozyme by Western blotting can only be attributed to a presumably fairly low occurrence of this protein (cf. different activity of the double promoter from cell to cell, possible instability of plant-produced lysozyme) or by insolubility that cannot be suppressed even by detergents. A correlation of the observed signal intensities in immunogold labeling with promoter activities cannot be carried out from the present data.

By distribution of lysozyme in the entire plant through transport in the phloem, dilution of local concentrations can additionally occur. On the other hand, a stimulation of the T_R promoter by such transport in the weakly expressing regions of the plant based on cell wall damage is supported.

The regeneration of plants that have integrated the plasmid pSR 4-3 with the two chimeric genes for the light and heavy chain of a monoclonal antibody successfully passed through the leaf disk test. During analysis of calli, which had been induced from leaf material from checked transformed plants, positive signals both for the light and heavy chain alone and for an intact antibody could be obtained by tissue printing (Cassab and Varner, 1987) from the total extract bound on the filter (detection of the light chain by polyclonal anti- λ and monoclonal anti- λ antibodies, of the heavy chain by polyclonal anti- μ antibodies and the aggregated B 1-8 by monoclonal antibodies that recognize an epitope in the tertiary structure from both chains close to the antigen bonding site). Since no compartments of the cell are separated here in an extraction pellet, but cell membranes and cell walls are together forced onto the filter, no distinction occurs concerning highly soluble extract, insoluble extract and organelle- or membrane-bound proteins and the spectrum of all proteins present in the cell is analyzed. If protein work-up is conducted by an affinity chromatographic method, a protein extract must first be produced. If only PBS is used as the extraction buffer, only a "soluble" extract is obtained. In order to solubilize bound or insoluble proteins as well, detergents must be added to the buffer, which mostly destroy the membranes and therefore liberate membrane-bound proteins or proteins enclosed by membranes (for example, in vesicles). For this purpose, Triton X-100 and CHAPS were used. Nevertheless, no significant improvements were observed. By extraction with SDS probe buffer, the presence of the light chain of the antibody could be demonstrated in plant tissue directly from the callus tissue, but was only reproducible with difficulty, i.e., generally only with weak signals. By isolation of the sought proteins with affinity gels from larger amounts of highly active leaf material, the yield should be significantly improved.

However, this does not apply to the light chain alone. Only a very weak signal could be obtained even by enrichment with a monoclonal antibody directed against the light chain of B 1-8. Detections of the heavy chain are not possible because of the more unfavorable conditions (weaker promoter and more insensitive detection system). Binding through the hapten is much milder than binding of B 1-8 through an anti-antibody and elution from the complex in a strongly acid medium. Since B 1-8 is a heteroclitic antibody, it recognizes the related hapten NIP better than NP and can consequently be eluted by NIP from a complex with NP. This occurs at neutral pH. In this manner, the existence of biologically active, intact B 1-8 in plant cells to be attained both from a soluble extract and from a CHAPS extract by incubation with NP Sepharose and subsequent detection of the light chain is demonstrated. A functional antibody that binds the hapten can only develop when correct assembly and correct folding of the light and heavy chains occur. Detection of the light chain after isolation by NP Sepharose is therefore an indirect but clear demonstration of the existence of intact B 1-8. No assertions can be made concerning the efficiency of the assembly because of the extremely limited amounts of protein. The question whether the heavy chain is glycosylated, as in lymphocytes, also cannot be answered with the available data. A functional antibody was measured in the ELISA test in *E. coli* after in vitro reassociation of genetic engineering-produced light and heavy chains from an extract even without glycosylation of a heavy chain (Boss et al., 1984), so that no inactivating factor is seen here. Possible effects on stability remain unaffected by this.

Only significantly weaker signals could be observed by immunogold labeling for the antibody analysis in comparison with T4 lysozyme. The Ac 38 antibody specifically recognizes only intact B 1-8 and can therefore be used for detection of a B 1-8 synthesized in plants. It could therefore be demonstrated that synthesis and assembling of this monoclonal antibody occurs on the rough endoplasmic reticulum. With the LS 136 antibody directed against the light chain, signals are found in the calli free in the cytoplasm, since both an isolated light chain and one aggregated together with the heavy chain can be involved and a clear assertion cannot be made from this type of experiment alone. By comparison with labelings with Ac 38, however, it is possible to conclude that the probability is greater that only a light chain is involved here. Since the double promoter is much more active in the callus than is the NOS promoter, it is conceivable that excess light chain that cannot be additionally glycosylated is removed from the secretion process and stored in the cytoplasm. Such a case for the precore protein of hepatitis B virus was recently described by Garcia et al. (1988), about 70 to 80% of this protein being localized in the cytoplasm in an in vitro experiment, despite cleavage of the signal peptide (i.e., from the ER). Processing by Western blotting also detected the B 1-8 antibody. Whereas an efficient secretion into the intercellular space could be demonstrated for the T4 lysozyme, there are still no available data for such an assertion concerning B 1-8, no labeling having yet been

found in the vicinity of the cell walls. However, it could not be ruled out that such labeling will still be detected during further more extensive investigations.

The relatively strong labelings in the chloroplasts occur surprisingly, since no structural properties are incorporated in the constructions that allow us to expect transport of the formed proteins into chloroplasts and the found signals involve signals for intact aggregated antibodies. However, since these labelings can be localized specifically in the chloroplasts, whereas the surroundings and control chloroplasts clearly have no labeling, a significance of these data must be assumed. Nevertheless, statistical solidification of this result must be waited for. It is structurally striking that the chloroplasts of the A5 plant contain unusual (electron-dense) regions. According to the now known view points, the signal peptide contains no features that should enable it to control transport into chloroplasts in addition to efficient secretion (α -amylase and T4 lysozyme). The question is therefore raised whether the amino acid sequences or the secondary or tertiary structures of the light or heavy chain or the aggregated antibody contain any information that is understood by the plants as a chloroplast transport signal. Should this phenomenon be clearly confirmed on additional sections, an additional field of highly interesting aspects would be opened up from these constructions.

The role of a BiP-like protein possibly present in the ER lumen (cf. Haas and Wabl, 1983; Bole et al., 1986; Hendershot et al., 1987) still remains open. In recent publications (Haas and Meo, 1988; Kassenbrock et al., 1988), however, it was reported that BiP is identical or related to other protein-binding proteins that were already previously identified. It could be shown that interactions occur not only with heavy chains in antibodies, but also with a larger number of other proteins. This supports the suspicion that such a class of proteins could also be present in plants and have a regulating role there of binding misfolded or incorrectly glycosylated proteins. Such an effect could also lead to the cytoplasmic localization of processed light chain as "reject."

The expression and assembling of a monoclonal antibody could be demonstrated by combination of Western blotting and immunogold labeling. It could be demonstrated by denaturing Western blotting that the light chain is processed as a result and, as supported by localization in the electron microscope, the foreign protein is inserted by the action of the α -amylase signal peptide into the lumen of the endoplasmic reticulum and secretion initiated. Assembling of the light and heavy chain also occurs there.

The relative intensity of the signal, on the one hand, could be attributed to a different stability of the foreign proteins in plants and, on the other hand, to expression-regulating sequences within the coding sequences. It is known that κ chains require an enhancer located in an intron for effective expression and without this only weaker expression occurs. If λ chains are cloned without such an enhancer, it is not weaker expression but no expression at all that can be

established; this observation could be due to the presence of inhibiting sequences within the light chain gene (A. Iglesias, personal communication). Such properties could also be expressed in plants. Additional effects could be produced by glycosylation or nonglycosylation of the foreign proteins in the ER and Golgi apparatus.

The approaches described in this work for resistance construction can be pursued in a variety of ways. On the one hand, a system is available that will be checked in a very short time for bacterial resistances. With a positive outcome, it is thus far the only known possibility for being able to take protective measures against bacterial infections of plants. A number of additional tests must naturally also occur in a toxicological respect.

After generation of monoclonal antibodies, for example, against the coat proteins of viruses, the biological efficiency of a reaction can be tested as a protective measure against viral infection with attacking viruses. Since the generation of antibodies is possible not only against virus coat proteins but against a number of other antigens, this method can be very widely extended and applied to create artificial resistance and for manipulation of a number of cellular processes. The basic principle, however, should be worked out with respect to simplifications, since expression and assembling of a heterotetrameric protein from two different genes is subject to a large number of influences and the conditions for statistical assembling of two times two different chains must be met. A significant simplification would be the application of a protein that has all the necessary components of a functional antibody localized on one gene. Such a construction of a semisynthetic "one-chain antibody" was produced by A. Iglesias and G. Kohler (Freiburg) and tested successfully on the formation of an idiotypical activity (unpublished results). The formation of such a chimeric protein should occur in the plant much more simply if the assembly of two chains is to be necessary to form the active center, this involving only one homodimeric protein. Experiments for preparation of such a project have already been started in cooperation with A. Iglesias and G. Kohler.

The synthesis of optimal resistance genes also includes infection- or wound-specific activation of these genes from a minimal expression level in the healthy state. In order to be able to cover all possible mechanisms, a number of special promoters will certainly be required. According to the present state of knowledge, only the T_R double promoter analyzed in this work offers itself for wound-specific activation. The characteristic properties of this promoter appear to make it suitable for this. The gene constructs described here therefore represent an important step on the way to developing functional "resistance genes."

The secretion of foreign proteins (also of prokaryotic origin) into the intercellular space made possible by the fusion of "resistance proteins" with a plant signal peptide opens up an additional field of optimization. By specific resistance proteins, pests could already be defended against before penetration into the cells or before damage to the cells. Interesting aspects are also

obtained from basic research into the secretion mechanism. The observation that information that controls localization must be obviously coded in the α -amylase signal peptide offers a contribution to the disputed question whether signal peptides only ensure insertion into the membrane of the rough endoplasmic reticulum or if they also have additional specificity for later localization. The major advantage of foreign proteins produced by genetic engineering in organisms is the absence of these proteins in the wild type organism so that clear assertions can be made concerning localization by comparison with the wild type. The properties of the foreign protein must then naturally be considered. Assembling of complex foreign proteins and secretion under the control of a plant signal peptide open up a broad field of research approaches for artificial resistances in plants.

V. Summary

By fusion of genes for plant-foreign proteins to the signal peptide gene of α -amylase from barley, the possibility of artificial secretion of such chimeric proteins produced by genetic engineering in plants was investigated. This approach was conducted, on the one hand, with a prokaryotic one-chain protein, lysozyme from bacteriophage T4, and, on the other hand, with a complex protein, a monoclonal IgM antibody. The assembly of two light and two heavy chains each into a tetrameric molecule is necessary for the formation of functional antibodies.

The genes for the light and heavy chain of the antibody and for the T4 lysozyme were all fused with the genes of the signal peptide of α -amylase. Promoter-fusion gene-polyadenylation sequence cassettes for the two antibody genes were constructed in a plasmid and for the T4 lysozyme in another plasmid. These constructions were integrated by agrobacterial infection and in the lysozyme also by naked DNA transfer in *Nicotiana tabacum*.

To have appropriate expression conditions for the promoters (pT_R and pNOS) for protein analysis, promoter activity analyses had to be conducted (by the NPT test). The T_R double promoter exhibits differential expression in intact plants with an expression maximum in the region of the transition from the stem to the roots. Much higher expression was found in calli (1000-fold). An up to 170-fold rise in activity within 6 days is attainable in an in vitro system by wounding or by hormone induction. The in vivo induction of greenhouse plants occurs within a day. Similar results were found for the NOS promoter, but with smaller factors (up to 20-fold inducible, callus no stronger than induced plants). For analysis of transgenic plants, highly sensitive nontraditional detection methods were worked out. Protein detection occurred, although in extremely limited amounts, by Western blotting with a biotin-streptavidin system, sometimes combined with an affinity chromatographic purification method from the protein crude extract. The presence of nonglycosylated unprocessed and processed T4 lysozyme, as well as possibly a glycosylated form could be demonstrated. By isolation of the light chain with an anti- λ antibody

and the intact antibody by functional binding to its hapten, both the presence and activity of a monoclonal antibody in the transgenic plant cells were demonstrated. In each case processed light chain was detected.

Analysis of the localization of these chimeric proteins in the cell occurred in cooperation with Dr. Sigrun Hippe (Aachen) by electron microscopic immunogold labeling. Strongly concentrated labeling was mostly found for T4 lysozyme in the intercellular spaces. Other labeled regions are present in the cytoplasm (mostly on filamentary structures) and inclusions of T4 lysozyme in proplastids of the phloem probably in crystalline form were repeatedly found, as well as labeling in and in the vicinity of plasmodesmata, on the cell walls of phloem and also in chloroplasts. The stored intercellularly T4 lysozyme forms, possibly in conjunction with other proteins, filamentary, network-like structures that are significant and were not found in the controls. The signal peptide of α -amylase from barley is therefore capable of localizing a prokaryotic protein in tobacco cells with specific characteristics and secreting it efficiently into the intercellular space.

The synthesis and assembling of the monoclonal antibody could be demonstrated by reaction with an anti-antibody in the endoplasmic reticulum directed against the intact antibody.

Processing of the signal peptide, as well as the functionality of the formed antibody, follow from the Western blotting analyses. As a result, secretion is also initiated by the signal peptide in this construction. Additional labelings occur in the chloroplasts and for the light chain also in the cytoplasm. Gold particles, however, could still not be detected in the region of the cell walls.

VI. Literature

- Abel, P.P., Nelson, R.S., De, B., Hoffmann, N., Rogers, S.G., Fraley, R.T., Beachy, R.N. (1986)
Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene
Science **232**, 738 - 743
- Ahl, P., Gianazzi, S. (1982)
b-Protein as a constitutive component in highly (TMV) resistant interspecific hybrids of *Nicotiana glutinosa* x *Nicotiana debneyi*
Plant Sci. Lett. **26**, 173 - 181
- Alt-Moerbe, J., Rak, B., Schröder, J. (1986)
A 3.6-kbp. segment from the vir region of Ti plasmids contains genes responsible for border sequence-directed production of T-region circles in *E. coli*
EMBO J. **5**, 1129 - 1135
- Akiyoshi, D.E., Morris, R.O., Hinz, R., Mischke, B.S., Koruge, T., Garfinkel, D.J., Gordon, M.P., Nester, E.W. (1983)
Cytokinin / auxin balance in crown gall tumors is regulated by specific loci on T-DNA
Proc. Natl. Acad. Sci. USA **80**, 407 - 411
- Akiyoshi, D.E., Klee, H., Amasino, R.M., Nester, E.W., Gordon, M.P. (1984)
T-DNA of *Agrobacterium tumefaciens* encodes an enzyme of cytokinin biosynthesis
Proc. Natl. Acad. Sci. USA **81**, 5994 - 5998
- Amasino, R.M., Miller, C.O. (1982)
Hormonal control of tobacco crown gall tumor morphology
Plant Physiol. **69**, 389 - 392
- Ashby, A.M., Watson, M.D., Shaw, C.H. (1987)
A Ti-plasmid determined function is responsible for chemotaxis of *Agrobacterium tumefaciens* towards the plant wound product acetosyringone
FEMS Microbiol. Lett. **41**, 189 - 192
- Austen, B., Hermon-Taylor, J., Kaderbhai, M.A., Ridd, D.H. (1984)
Design and synthesis of a consensus signal sequence that inhibits protein translocation into rough microsomal vesicles
Biochem. J. **224**, 318 - 323
- Baulcombe, D.C., Saunders, G.R., Bevan, M.W. (1986)
Expression of biologically active viral satellite RNA in the genome of transformed plants
Nature **321**, 446 - 449
- Beckers, C.J.M., Keller, D.S., Balch, W.E. (1987)
Semi-intact cells permeable to macromolecules: protein transport from the endoplasmic reticulum
Cell **50**, 523 - 534
- Benson, S.A., Hall, M.N., Silhavy, T.J. (1985)
Genetic analysis of protein export in *Escherichia coli*
Ann. Rev. Biochem. **54**, 101 - 134
- Berger, S., de Groot, E.J., Neuhaus, G., Schweiger, M. (1987)
Acetabularia: a giant single cell organism with valuable advantages for cell biology
Eur. J. Cell Biol. **44**, 349 - 370
- Bernal-Lugo, I., Beachy, R.N., Varner, J.E. (1981)
The response of barley aleurone layers to gibberellic acid includes the transcription of new sequences
Biochem. Biophys. Res. Comm. **102**, 617 - 623

- Bernasconi, P., Jollès, P., Pilet, P.E. (1986)
Increase of lysozyme and chitinase in *Rubus* caused by infection and some polymers
Plant Sci. 44, 79 - 83
- Bernier, I. (1971)
The turnip lysozyme
FEBS Lett. 14, 100 - 104
- Bevan, M., Barnes, W.M., Chilton, M.-D. (1983)
Structure and transcription of the nopaline synthase gene region of T-DNA
Nucl. Acids Res. 11, 369 - 385
- Bevan, M. (1984)
Binary *Agrobacterium* vectors for plant transformation
Nucl. Acids Res. 12, 8711 - 8721
- Bevan, M.W., Mason, S.E., Goelet, P. (1985)
Expression of tobacco mosaic virus coat protein by a cauliflower mosaic virus promoter in plants transformed by *Agrobacterium*
EMBO J. 4, 1921 - 1926
- Berg, P., Sambrook, J. (1987)
In yeast and mammalian cells: not all signal sequences are equivalent
Cell 50, 2903 - 2914
- Doly, J. (1979)
A simple extraction procedure for screening recombinant plasmid
Nucl. Acids Res. 7, 1513 - 1525
- Eickbush, D., Eickbush, M., Strub, K. (1985)
Poly(A) termination and 3' processing: The end is in site
Nucl. Acids Res. 13, 349 - 359
- Blobel, G., Dobberstein, B. (1975a)
Transfer of proteins across membranes
I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma
J. Cell Biol. 67, 835 - 851
- Blobel, G., Dobberstein, B. (1975b)
Transfer of proteins across membranes
II. Reconstitution of functional rough microsomes from heterologous compounds
J. Cell Biol. 67, 852 - 862
- Bole, D.G., Hendershot, L., Kearney, J.F. (1986)
Posttranslational association of immunoglobulin heavy chain binding protein with nascent heavy chains in nonsecreting and secreting hybridomas
J. Cell Biol. 102, 1558 - 1566
- Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W. (1977)
Construction and characterization of new cloning vehicles II: A multipurpose cloning system
Gene 2, 95 - 113
- Boller, T., Gehri, A., Mauch, F., Vögeli, U. (1983)
Chitinase in bean leaves: induction by ethylene, purification, properties and possible function
Planta 157, 22 - 31

- Bolton, G.W., Nester, E.W., Gordon, M.P. (1986)
Plant phenolic compounds induce expression of the *Agrobacterium tumefaciens* loci needed for virulence
Science 232, 983 - 985
- Boss, M.A., Kenten, J.H., Wood, C.R., Emtage, J.S. (1984)
Assembly of functional antibodies from immunoglobulin heavy and light chains synthesized in *E.coli*
Nucl.Acids Res. 12, 3791 - 3806
- Boston, R.S., Becwar, M.R., Ryan, R.D., Goldsbrough, P.B., Larkins, B.A., Hodges, T.K. (1987)
Expression from heterologous promoters in electroporated carrot protoplasts
Plant Physiol. 83, 742 - 746
- Bothwell, A.L.M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewski, K., Baltimore, D. (1981)
Heavy chain variable region contribution to the NP^b family of antibodies: somatic mutation evident in a $\gamma 2a$ variable region
Cell 24, 625 - 637
- Bothwell, A.L.M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewski, K., Baltimore, D. (1982)
Somatic variants of murine immunoglobulin light chains
Nature 298, 380 - 382
- Bradford, M.M. (1976)
A rapid and sensitive method for quantification of microgram quantities of protein utilizing the principle of protein-dye binding
Anal.Biochem. 72, 248 - 254
- Briggs, M.S., Cornell, D.G., Dluhy, R.A., Gierasch, L.M. (1986)
Conformations of signal peptides induced by lipids suggest initial steps in protein export
Science 233, 206 - 208
- Broadway, R.M., Duffey, S.S., Pearce, G., Ryan, C.A. (1986)
Plant proteinase inhibitors: a defense against herbivorous insects
Entomol.exp.appl. 41, 33 - 38
- Buchanan-Wollaston, V., Passiatore, J.E., Cannon, F. (1987)
The mob and ori T mobilization functions of a bacterial plasmid promote its transfer to plants
Nature 328, 172 - 175
- Budar, F., Deboeck, F., Van Montagu, M., Hernalsteens, J.-P. (1986)
Introduction and expression of the octopine T-DNA oncogenes in tobacco plants and their progeny
Plant Sci. 46, 195 - 206
- Burnette, W.N. (1981)
"Western Blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate - polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radiolabeled Protein A
Anal.Biochem. 112, 195 - 203
- Cabilly, S., Riggs, A.D., Pande, H., Shively, J.E., Holmes, W.E., Roy, M., Perry, L.J., Wetzel, R., Heyneker, H.L. (1984)
Generation of antibody activity from immunoglobulin polypeptide chains produced in *E.coli*
Proc.Natl.Acad.Sci. USA 81, 3273 - 3277
- Caplan, A.B., Van Montagu, M., Schell, J. (1985)
Genetic analysis of integration mediated by single T-DNA borders
J.Bacteriol. 161, 655 - 664
- Carlén, E., Kellenberger, E. (1982)
The reproducible observation of unstained embedded cellular material in thin sections: visualization of an integral membrane protein by a new mode of imaging for STEM
EMBO J. 1, 63 - 67

- Carr, J.P., Dixon, D.C., Nikolau, B.J., Voelkerding, K.V., Kleissig, D.F. (1987)
Synthesis and localization of pathogenesis-related proteins in tobacco
Mol. Cell Biol. 7, 1580 - 1583
- Cassab, G.I., Varner, J. (1987)
Immunocytochemical localization of extensin in developing soybean seed coats by immunogold-silver staining and by tissue printing on nitrocellulose paper
J. Cell Biol. 105, 2581 - 2588
- Chaiet, L., Wolf, F.J. (1964)
The properties of streptavidin, a biotin-binding protein produced by *Streptomyces*
Arch. Biochem. Biophys. 106, 1 - 5
- Chan, V.T.-W., Fleming, K.A., McGee, J.O'D. (1985)
Detection of subpicogram quantities of specific DNA sequences on blot hybridization with biotinylated probes
Nucl. Acids Res. 13, 8083 - 8091
- Chandler, P.M., Zwar, J.A., Jacobsen, J.V., Higgins, T.J.V., Inglis, A.S. (1984)
The effect of gibberellic acid and abscisic acid on α -amylase mRNA levels in barley aleurone layers studies using an α -amylase cDNA clone
Plant Mol. Biol. 3, 407 - 418
- Chen, J., Varner, J.E. (1985a)
Isolation and characterization of cDNA clones for carrot extensin and a proline-rich 33-kDa protein
Proc. Natl. Acad. Sci. USA 82, 4399 - 4403
- Chen, J., Varner, J.E. (1985b)
An extracellular matrix protein in plants: characterization of a genomic clone for carrot extensin
EMBO J. 4, 2145 - 2151
- Coleman, J., Inukai, M., Inouye, M. (1985)
Dual functions of the signal peptide in protein transfer across the membrane
Cell 43, 351 - 360
- Cornai, L., Facchetti, D., Hiatt, N.R., Thompson, G., Rose, R.E., Stalker, D.M. (1985)
Expression in plants of a mutant *aroA* gene from *Salmonella typhimurium* confers tolerance to glyphosate
Nature 317, 741 - 744
- Cornelissen, B.J.C., Horowitz, J., van Kan, J.A.L., Goldberg, R.B., Bol, J.F. (1987)
Structure of tobacco genes encoding pathogenesis-related proteins from the PR-1 group
Nucl. Acids Res. 15, 6799 - 6811
- Courtice, G. (1987)
Satellite defense for plants
Nature 328, 758 - 759
- Craig, S., Staehelin, A. (1988)
High pressure freezing of intact plant tissues. Evaluation and characterization of novel features of the endoplasmic reticulum and associated membrane systems
Eur. J. Cell Biol. 46, 80 - 93
- Czernilofsky, A.P., Hain, R., Herrera-Estrella, L., Lötz, H., Goyvaerts, E., Baker, B.J., Schell, J. (1986)
Fate of selectable marker DNA integrated into the genome of *Nicotiana tabacum*
DNA 5, 101 - 113
- Dagert, M., Ehrlich, S.D. (1979)
Prolonged incubation in calcium chloride improves the competence of *E. coli* cells
Gene 6, 23 - 28
- Darvill, A.G., Albersheim, P. (1984)
Phytoalexins and their elicitors - A defense against microbial infection in plants
Ann. Rev. Plant Physiol. 35, 243 - 275

- Davis, K.R., Darvill, A.G., Albersheim, P. (1986)
Several biotic and abiotic elicitors act synergistically in the induction of phytoalexin accumulation in soybean
Plant Mol.Biol. 6, 23 - 32
- Dean, R.A., Kuc, J. (1985)
Induced systemic protection in plants
Trends Biotechnol. 3, 125 - 129
- De Block, M., Herrera-Estrella, L., Van Montagu, M., Schell, J. (1984)
Expression of foreign genes in regenerated plants and in their progeny
EMBO J. 3, 1681 - 1689
- De Block, M., Botterman, J., Vandewiele, M., Dockx, J., Thoen, C., Gessel, V., Rao Movva, N., Thompson, C., Van Montagu, M., Leemans, J. (1987)
Engineering herbicide resistance in plants by expression of a detoxifying enzyme
EMBO J. 6, 2513 - 2518
- Depicker, A., Stachel, S., Dhaese, P., Zambryski, P., Goodman, H.M. (1982)
Nopaline synthase: transcript mapping and DNA sequence
J.Mol.Appl.Genet. 1, 561 - 573
- Depicker, A., Herman, L., Jacobs, A., Schell, J., Van Montagu, M. (1985)
Frequencies of simultaneous transformation with different T-DNAs and their relevance to the *Agrobacterium* plant cell
Mol.Gen.Genet. 201, 477 - 484
- Di Rita, V.J., Gelvin, S.B. (1987)
Deletion analysis of the mannopine synthase gene promoter in sunflower crown gall tumors and *Agrobacterium tumefaciens*
Mol.Gen.Genet. 207, 233 - 241
- Donn, G., Fischer, E., Smith, J.A., Goodman, H.M. (1984)
Herbicide-resistant alfalfa cells: an example of gene amplification in plants
J.Mol.Appl.Genet. 2, 621 - 635
- Dörner, A.J., Bole, D.G., Kaufman, R.J. (1987)
The relationship of N-linked glycosylation and heavy chain - binding protein association with the secretion of glycoproteins
J.Cell Biol. 105, 2665 - 2674
- Douglas, C., Staneloni, R.J., Rubin, R.A., Nester, E.W. (1985)
Identification and genetic analysis of an *Agrobacterium tumefaciens* chromosomal virulence region
J.Bacteriol. 161, 850 - 860
- Duong, L.T., Caulfield, M.P., Rosenblatt, M. (1987)
Synthetic signal peptide and analogs display different activities in mammalian and plant in vitro secretion systems
J.Biol.Chem. 262, 6328 - 6333
- Ebert, P.R., Ha, S.B., An, G. (1987)
Identification of an essential upstream element in the nopaline synthase promoter by stable and transient assay
Proc.Natl.Acad.Sci.USA 84, 5745 - 5749
- Ecker, P., Donn, G., Wengenmayer, F. (1987)
Genetic engineering with plants.
Angew. Chem. 99, 392-412.
- Ellis, J.G., Ryder, M.H., Tate, M.E. (1984)
Agrobacterium tumefaciens T₂ - DNA encodes a pathway for agropine biosynthesis
Mol.Gen.Genet. 195, 466 - 473
- Engström, P., Zambryski, P., Van Montagu, M., Stachel, S. (1987)
Characterization of *Agrobacterium tumefaciens* virulence proteins induced by the plant factor acetosyringone
J.Mol.Biol. 197, 635 - 645

- **Feinberg, A.P., Vogelstein, B. (1983, 1984)**
A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity
Anal. Biochem. 132, 6 - 13
Anal. Biochem. 137, 266 - 267
- **Fischhoff, D.A., Bowditch, K.S., Perlak, F.J., Marrone, P.G., McCormick, S.M., Niedermeyer, J.G., Dean, D.A., Kussano - Kretzmer, K., Mayer, E.J., Rochester, D.E., Rogers, S.G., Fraley, R.T. (1987)**
Insect tolerant transgenic tomato plants
Biotechnol. 5, 801 - 813
- **Fernandez, D.E., Staehelin, L.A. (1985)**
Structural organization of ultrarapidly frozen barley aleurone cells actively involved in protein secretion
Planta 165, 455 - 468
- **Ferro-Novick, S. (1985)**
Genetic approaches to the study of protein targeting across a lipid bilayer
Trends Biochem. Sci. 10, 425 - 427
- **Fromm, M., Taylor, L.P., Walbot, V. (1985)**
Expression of genes transferred into monocot and dicot plant cells by electroporation
Proc. Natl. Acad. Sci. USA 82, 5824 - 5828
- **Fromm, M., Walbot, V. (1987)**
Transient expression of DNA in plant cells
in: *Plant Gene Research, Plant DNA Infectious Agents*, ed.: Hohn, T., Schell, J., pp. 303 - 310
Springer Verlag Wien, New York
- **Garcia, P.D., Ou, J.-H., Rutter, W.J., Walter, P. (1988)**
Targeting of the hepatitis B virus precore protein to the endoplasmic reticulum membrane: after signal peptide cleavage translocation can be aborted and the product released into the cytoplasm
J. Cell Biol. 106, 1093 - 1104
- **Gardner, R.C., Knäuf, V.C. (1986)**
Transfer of *Agrobacterium* DNA to plants requires a T-DNA border but not the *virE* locus
Science 231, 725 - 727
- **Gelvin, S.B., Gordon, M.P., Nester, E.W., Aronson, A.J. (1981)**
Transcription of the *Agrobacterium* Ti plasmid in the bacterium and in crown gall tumors
Plasmid 6, 17 - 29
- **Gelvin, S.B., Karcher, S.J., Goldsbrough, P.B. (1985)**
Use of a T₂ T-DNA promoter to express genes in plants and bacteria
Mol. Gen. Genet. 199, 240 - 248
- **Gerlach, W.L., Llewellyn, D., Hasehoff, J. (1987)**
Construction of a plant disease resistance gene from the satellite RNA of tobacco ringspot virus
Nature 328, 802 - 805
- **Gershoni, J.M., Palade, G.E. (1983)**
Protein blotting: principles and applications
Anal. Biochem. 131, 1 - 15
- **Gibson, R., Schlesinger, S., Kornfeld, S. (1979)**
The non-glycosylated glycoprotein of vesicular stomatitis virus is temperature-sensitive and undergoes intracellular aggregation at elevated temperatures
J. Biol. Chem. 254, 3600 - 3607
- **Gilmore, R., Blobel, G. (1983)**
Transient involvement of signal recognition particle and its receptor in the microsomal membrane prior to protein translocation
Cell 35, 677 - 685

- Gilmore, R., Blobel, G. (1985)
Translocation of secretory proteins across the microsomal membrane occurs through an environment accessible to aqueous perturbants
Cell 42, 497 - 505
- Glazer, A.N., Barel, A.O., Howard, J.B., Brown, D.M. (1969)
Isolation and characterization of fig lysozyme
J.Biol.Chem. 244, 3583 - 3589
- Green, N.M., Toms, E.J. (1970)
Purification and crystallization of avidin
Biochem.J. 118, 67 - 70
- Griffiths, G., Simons, K. (1986)
The trans Golgi network: sorting at the exit site of the Golgi complex
Science 234, 438 - 443
- Grütter, M.G., Weaver, L.H., Gray, T.M., Matthews, B.W. (1983)
Structure, function and evolution of the lysozyme from bacteriophage T4
in: *Bacteriophage T4*,
ed.: Mathews, C.K., Kutter, E.M., Mosig, G., Bergel, P.B., pp. 356 - 360
American Society for Microbiology, Washington, D.C.
- Guan, J.-L., Machamer, C.E., Ross, J.K. (1985)
Glycosylation allows cell surface transport of an anchored secretory protein
Cell 42, 489 - 496
- Gubler, F., Jacobsen, J.V., Ashford, A.E. (1986)
Involvement of the Golgi apparatus in the secretion of α -amylase from gibberellin-treated barley aleurone cells
Planta 168, 447 - 452
- Gubler, F., Ashford, A.E., Jacobsen, J.V. (1987)
The release of α -amylase through gibberellin-treated aleurone cell walls
Planta 172, 155 - 161
- Haas, I.G., Wabl, M. (1983)
Immunoglobulin heavy chain binding protein
Nature 306, 387 - 389
- Haas, I.G., Mee, T. (1988)
cDNA cloning of the immunoglobulin heavy chain binding protein
Proc.Natl.Acad.Sci.USA 85, 2230 - 2234
- Haerdtle, M.-T., Aubert, M.L., Djiane, J., Kraehenbuhl, J.-P. (1983)
Binding sites for lactogenic and somatogenic hormones from rabbit mammary gland and liver
J.Biol.Chem. 258, 305 - 314
- Hain, R., Stabel, P., Czernilofsky, A.P., Steinbis, H.H., Herrera-Estrella, L., Schell, J. (1985)
Uptake, integration, expression and genetic transmission of a selectable chimeric gene by plant protoplasts
Mol.Gen.Genet. 199, 161 - 168
- Hanahan, D. (1983)
Studies on transformation of *Escherichia coli* with plasmids
J.Mol.Biol. 166, 557 - 580
- Harpster, M.H., Townsend, J.A., Jones, J.D.G., Bedbrook, J., Dunsmuir, P. (1988)
Relative strengths of the 35S cauliflower mosaic virus, 1', 2', and nopaline synthase promoters in transformed tobacco, sugarbeet and oilseed rape callus tissue
Mol.Gen.Genet. 212, 182 - 190
- Harrison, B.D., Mayo, M.A., Baulcombe, D.C. (1987)
Virus resistance in transgenic plants that express cucumber mosaic virus satellite RNA
Nature 328, 799 - 802
- Heidecker, G., Messing, J. (1986)
Structural analysis of plant genes
Ann.Rev.Plant Physiol. 37, 439 - 466

- Hein, I., Preszowozny, I., Schlieder, O. (1983)
Culture and selection of somatic hybrids using an auxotrophic cell line
Theor. Appl. Genet. 64, 119 - 122
- Hendershot, L., Bole, D., Köhler, G., Kearney, J.P. (1987)
Assembly and secretion of heavy chains that do not associate
posttranslationally with immunoglobulin heavy chain - binding protein
J. Cell Biol. 104, 761 - 767
- Hepburn, A.G., Clarke, L.E., Pearson, L., White, J. (1983)
The role of cytosin methylation in the control of nopaline synthase gene
expression in a plant tumor
J. Mol. Appl. Genet. 2, 315 - 329
- Hernalsteens, J.-P., Thia-Taong, L., Schell, J., Van Montagu, M. (1984)
An Agrobacterium transformed cell culture from the monocot *Asparagus
officinalis*
EMBO J. 3, 3039 - 3041
- Higgins, T.J.V., Jacobsen, J.V., Zwar, J.A. (1982)
Gibberellic acid and abscisic acid modulate protein synthesis and mRNA
levels in barley aleurone layers
Plant Mol. Biol. 1, 191 - 215
- Hilder, V.A., Gatehouse, A.M.R., Sheerman, S.E., Barker, R.F., Boulter, D.
(1987)
A novel mechanism of insect resistance engineered in tobacco
Nature 330, 160 - 163
- Hippe, S., Hermanns, M. (1986)
Improved structural preservation in freeze-substituted sporidia of *Ustilago
avenae* - a comparison with low-temperature embedding
Protoplasma 135, 19 - 30
- Hippe, S. (1987)
Combined application of low temperature preparation and electron
microscopic autoradiography for the localization of systemic fungicides
Histochem. 87, 309 - 315
- Hoekema, A., de Pater, B.S., Fellingner, A.J., Hooykaas, P.J.J.,
Schilperoort, R.A.A. (1984)
The limited host range of an *Agrobacterium tumefaciens* strain extended
by a cytokinin gene from a wide host range T-region
EMBO J. 3, 3043 - 3047
- Hoekema, A., Van Haaren, M.J.J., Fellingner, A.J., Hooykaas, P.J.J.,
Schilperoort, R.A. (1985)
Non-oncogenic plant vectors for use in the *agrobacterium* binary system
Plant Mol. Biol. 5, 85 - 89
- Hoffmann, K., Wood, S.W., Brinton, C.C., Montibeller, J.A., Finn, F.M. (1980)
Immunoblot affinity columns and their application to retrieval of
streptavidin
Proc. Natl. Acad. Sci. USA 77, 4666 - 4668
- Hooft van Hujsduijnen, R.A.M., Cornelissen, B.J.C., Van Loon, L.C.,
Van Boom, J.H., Tromp, M., Bol, J.F. (1985)
Virus-induced synthesis of messenger RNAs for precursors of
pathogenesis-related proteins in tobacco
EMBO J. 4, 2167 - 2171
- Hooft van Hujsduijnen, R.A.M., Van Loon, L.C., Bol, J.F. (1986)
cDNA cloning of six mRNAs induced by TMV infection of tobacco and a
characterization of their translation products
EMBO J. 5, 2057 - 2061
- Hooykaas, P.J.J., Schilperoort, R.A. (1985)
The Ti-plasmid of *Agrobacterium tumefaciens*: a natural genetic engineer
Trends Biotechnol. Sci. 307 - 309
- Horsch, R.B., Fraley, R.T., Rogers, S.G., Sanders, P.R., Lloyd, A., Hoffmann, N.
(1984)
Inheritance of functional foreign genes in plants
Science 223, 496 - 498

- Horsch,R.B., Fry,J.E., Hoffmann,N.L., Eichholtz,D., Rogers,S.G., Fraley,R.T. (1985)
A simple and general method for transferring genes into plants
Science 227, 1229 - 1231
- Horsch,R.B., Klee,H.J. (1986)
Rapid assay of foreign gene expression in leaf discs transformed by *Agrobacterium tumefaciens*: role of T-DNA borders in the transfer process
Proc.Natl.Acad.Sci.USA 83, 4428 - 4432
- Hortsch,M., Avossa,D., Meyer,D.I. (1985)
A structural and functional analysis of the docking protein
J.Biol.Chem. 260, 9137 - 9145
- Howard,J.B., Glazer,A.N. (1967)
Studies on the physicochemical and enzymatic properties of Papaya lysozyme
J.Biol.Chem. 242, 5717 - 5723
- Huang,J.-K., Swegle,M., Dandekar,A.M., Muthukrishnan,S. (1984)
Expression and regulation of α -amylase gene family in barley aleurones
J.Mol.Appl.Genet. 2, 579 - 588
- Jacobsen,J.V., Knox,R.B. (1973)
Cytochemical localization and antigenicity of α -amylase in barley aleurone tissue
Planta 112, 213 - 224
- Janssens,A., Genetello,C., Van Montagu,M., Zambryski,P. (1986)
Plant cells induce transcription of the *Agrobacterium tumefaciens* nopaline pTi C58 virulence region
Plant Sci. 47, 185 - 193
- Jones,J.D.G., Dunsmuir,P., Bedbrook,J. (1985)
High level expression of introduced chimeric genes in regenerated transformed plants
EMBO J. 4, 2411 - 2418
- Jones,J.D.G., Grady,K.L., Suslow,T.V., Bedbrook,J.R. (1986)
Isolation and characterization of genes encoding two chitinase enzymes from *Serratia marcescens*
EMBO J. 5, 467 - 473
- Jones,J.D.G., Svab,Z., Harper,R.C., Hurwitz,C.D., Maliga,P. (1987)
A dominant nuclear streptomycin resistance marker for plant cell transformation
Mol.Gen.Genet. 210, 86 - 91
- Jones,R.L., Chen,R.F. (1976)
Immunohistochemical localization of α -amylase in barley aleurone cells
J.Cell Sci. 20, 183 - 198
- Jones,R.L., Jacobsen,J.V. (1982)
The role of the endoplasmic reticulum in the synthesis and transport of α -amylase in barley aleurone layers
Planta 156, 421 - 432
- Joshi,C.P. (1987)
An inspection of the domain between putative TATA box and translation start site in 79 plant genes
Nucl.Acids Res. 15, 6643 - 6653
- Josslin,R. (1970)
The lysis mechanism of phage T4: mutants affecting lysis
Virology 40, 719 - 726
- Kaiser,C.A., Preuss,D., Grisafi,P., Botstein,D. (1987)
Many random sequences functionally replace the secretion signal sequence of yeast invertase
Science 235, 312 - 317

- Kao, K.N., Constabel, F., Michayluk, M.R., Gamborg, O.L. (1974)
Plant protoplast fusion and growth of intergeneric hybrid cells
Planta 120, 215 - 227
- Karcher, S.J., Di Rita, V.J., Gelvin, S.B. (1984)
Transcript analysis of T₂-DNA in octopine-type crown gall tumors
Mol.Gen.Genet. 194, 159 - 165
- Kassenbrock, C.K., Garcia, P.D., Walter, P., Kelly, R.B. (1988)
Heavy-chain binding protein recognizes aberrant polypeptides translocated in vitro
Nature 333, 90 - 93
- Kauffmann, S., Legrand, M., Geoffroy, P., Fritig, B. (1987)
Biological function of "pathogenesis-related" proteins: four proteins of tobacco have 1,3- β -glucanase activity
EMBO J. 6, 3209 - 3212
- Kaulen, H. (1986)
Studies on light-induced expression of a chimeric chalcone synthase neomycin phosphotransferase II gene in *Nicotiana tabacum*.
Dissertation, Cologne University
- Kay, R., Chan, A., McPherson, J. (1987)
Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes
Science 236, 1299 - 1302
- Kellenberger, E., Carlén, E., Villiger, W., Roth, J., Garavito, R.M. (1980)
Low denaturation embedding for electron microscopy of thin sections.
Publiziert von: Chemische Werke Lowi GmbH, S. 1 - 59
- Kelly, R.B. (1985)
Pathways of protein secretion in eucaryotes
Science 230, 25 - 32
- Kemper, E., Waffenschmidt, S., Weller, E.W., Schröder, J. (1985)
T-DNA-encoded auxin formation in crown-gall cells
Planta 163, 257 - 262
- Kendall, D.A., Bock, S.C., Kaiser, E.T. (1986)
Idealization of the hydrophobic segment of the alkaline phosphatase signal peptide
Nature 321, 706 - 708
- Kombrink, E., Schröder, M., Hahlbrock, K. (1988)
Several "pathogenesis-related" proteins in potato are β -1,3-glucanases and chitinases
Proc.Natl.Acad.Sci.USA 85, 782 - 786
- Koncz, C., De Greve, H., André, D., Deboeck, F., Van Montagu, M., Schell, J. (1983)
The opine synthase genes carried by Ti plasmids contain all signals necessary for expression in plants
EMBO J. 2, 1597 - 1603
- Koncz, C., Schell, J. (1986)
The promoter of T₂-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector
Mol.Gen.Genet. 204, 383 - 396
- Koncz, C., Olsson, O., Langridge, W.H.R., Schell, J. (1987)
Expression and assembly of functional bacterial luciferase in plants
Proc.Natl.Acad.Sci.USA 84, 131 - 135
- Koukolikova-Nicola, Z., Shillito, R.D., Hohn, B., Wang, K., Van Montagu, M., Zambryski, P. (1985)
Involvement of circular intermediates in the transfer of T-DNA from *Agrobacterium tumefaciens* to plant cells
Nature 313, 191 - 196

- Krens, F.A., Molendijk, L., Willems, G.J., Schilperoort, R.A. (1982)
In vitro transformation of plant protoplasts with Ti-plasmid DNA
Nature 296, 72 - 74
- Kuhn, A., Wickner, W. (1985)
Conserved residues of the leader peptide are essential for cleavage by leader peptidase
J.Biol.Chem. 260, 15914 - 15918
- Kuhn, A., Wickner, W., Kreil, G. (1986)
The cytoplasmic carboxy terminus of M13 procoat is required for the membrane insertion of its central domain
Nature 322, 335 - 339
- Kurzchalia, T.V., Wiedmann, M., Ginzovich, A.S., Bochkareva, E.S., Bielka, H., Rapoport, T.A. (1986)
The signal sequence of nascent preprolactin interacts with the 54K polypeptide of the signal recognition particle
Nature 320, 634 - 636
- Kyhse-Andersen, J. (1984)
Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose
J.Biochem.Biophys.Meth. 10, 203
- Kozak, M. (1986)
Point mutations define a sequence flanking the AUG codon that modulates translation by eucaryotic ribosomes
Cell 44, 283 - 292
- Laemmli, U.K. (1970)
Cleavage of structural proteins during the assembly of the head of bacteriophage T4
Nature 227, 680 - 685
- Langer, P.R., Waldrop, A.A., Ward, D.C. (1981)
Enzymatic synthesis of biotin-labeled polynucleotides: Novel nucleic acid affinity probes
Proc.Natl.Acad.Sci.USA 78, 6633 - 6637
- Langridge, P., Brown, J.W.S., Pinter-Toro, J.A., Feix, G., Neuhaus, G., Neuhaus-Uri, G., Schweiger, H.G. (1985)
Expression of zein genes in *Acetabularia mediterranea*
Eur.J.Cell Biol. 39, 257 - 264
- Lauffer, L., Garcia, P.D., Harkins, R.N., Coussens, L., Ulrich, A., Walter, P. (1985)
Topology of signal recognition particle receptor in endoplasmic reticulum membrane
Nature 318, 334 - 338
- Leary, J.J., Brigati, D.J., Ward, D.C. (1983)
Rapid and sensitive colorimetric method for visualizing biotin-labeled DNA probes hybridized to DNA or RNA immobilized on nitrocellulose: Bio-blot
Proc.Natl.Acad.Sci.USA 80, 4045 - 4049
- Lee, J.S., Brown, W.E., Graham, J.S., Pearce, G., Fox, E.A., Dreher, T.W., Ahren, K.G., Pearson, G.D., Ryan, A. (1986)
Molecular characterization and phylogenetic studies of a wound-inducible proteinase inhibitor I gene in *Lycopersicon* species
Proc.Natl.Acad.Sci.USA 83, 7277 - 7281
- Leemans, J., Deblaeze, R., Willmitzer, L., De Greve, H., Hernalsteens, J.-P., Van Montagu, M., Schell, J. (1982)
Genetic identification of functions of T_R - DNA transcripts in octopine crown galls
EMBO J. 1, 147 - 152
- Legrand, M., Kauffmann, S., Geoffroy, P., Fritig, B. (1987)
Biological function of pathogenesis-related proteins: Four tobacco pathogenesis-related proteins are chitinases
Proc.Natl.Acad.Sci.USA 84, 6750 - 6754
- Linsmaier, E.M., Skoog, F. (1965)
Organic growth factor requirements of tobacco tissue cultures
Physiol. Plant. 18, 100 - 127

- Looy, R., Kende, H. (1978)
The mode of secretion of α -amylase in barley aleurone layers
Planta 143, 89 - 99
- Lodish, H.F., Kong, N., Hirani, S., Rasmussen, J. (1987)
A vesicular intermediate in the transport of hepatoma secretory proteins from the rough endoplasmic reticulum to the Golgi complex
J. Cell Biol. 104, 221 - 230
- Lodish, H.F. (1988)
Transport of secretory and membrane glycoproteins from the rough endoplasmic reticulum to the Golgi
J. Biol. Chem. 263, 2107 - 2110
- Logemann, J., Schell, J., Willmitzer, L. (1987)
Improved method for the isolation of RNA from plant tissues
Anal. Biochem. 163, 16 - 20
- Low, B. (1968)
Formation of merodiploids in matings with a class of Rec⁻ recipient strains of *E. coli* K12
Proc. Natl. Acad. Sci. USA 60, 160 - 167
- Machida, Y., Usami, S., Yamamoto, A., Niwa, Y., Takebe, I. (1986)
Plant-inducible recombination between the 25bp border sequences of T-DNA in *Agrobacterium tumefaciens*
Mol. Gen. Genet. 204, 374 - 382
- Maliga, P., Sz. Breznovitis, A., Marton, L. (1973)
Streptomycin-resistant plants from callus culture of haploid tobacco
Nature New Biol. 244, 29 - 30
- Maniatis, T., Fritsch, E.F., Sambrook, J. (1982)
Molecular cloning, a laboratory manual
Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Matthews, B.W., Remington, S.J., Gruetter, M.G., Anderson, W.F. (1981)
Relation between hen egg white lysozyme and bacteriophage T4 lysozyme: evolutionary implications
J. Mol. Biol. 147, 545 - 558
- Matthysse, A.G. (1986)
Effect of the presence of the plasmid pSA and of auxin on the attachment of *Agrobacterium tumefaciens* to the plant host cells
In: *Mol. Genet. Plant-Microbe Interactions, Proc. Int. Symp. 3. Meeting.* S. 11 - 13
ed.: Verma, D.P.S., Brisson, N.
Nijhoff, Dordrecht, Nederlande
- Matze, A.J.M., Chilton, M.-D. (1981)
Site-specific insertion of genes into T-DNA of the *Agrobacterium* tumor-inducing plasmid: an approach to genetic engineering of higher plant cells
J. Mol. Appl. Genet. 1, 39 - 49
- McKenzie, H.A., White, Jr., F.H. (1986)
Determination of lysozyme activity at low levels with emphasis on the milk enzyme
Anal. Biochem. 157, 367 - 374
- McPheeters, D.S., Christensen, A., Young, E.T., Stormo, G., Gold, L. (1986)
Translational regulation of the bacteriophage T4 lysozyme gene
Nucl. Acids Res. 14, 5813 - 5826
- Melchers, L.S., Thompson, D.V., Idler, K.B., Neuteboom, S.T.C., de Maagd, R.A., Schilperoort, R.A., Hooykaas, P.J.J. (1987)
Molecular characterization of the virulence gene *virA* of *Agrobacterium tumefaciens* octopine T1 plasmid
Plant Mol. Biol. 9, 635 - 645
- Meyer, D.I., Krause, E., Dobberstein, B. (1982)
Secretory protein translocation across membranes - the role of the "docking protein"
Nature 297, 647 - 650
- Miller, J.H. (1972)
Experiments in molecular genetics
Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

- Mize,N.K., Andrews,D.W., Lingappa,V.R. (1986)
A stop transfer sequence recognizes receptors for nascent chain translocation across the endoplasmic reticulum membrane
Cell 47, 711 - 719
- Moor,H. (1987)
Theory and Practice of high pressure freezing
in: Cryotechniques in Biological Electron Microscopy
ed.: Steinbrecht,R.A., Zierold,K., S. 175 - 191
Springer - Verlag Berlin, Heidelberg
- Morris,R.O. (1986)
Genes specifying auxin and cytokinin biosynthesis in phytopathogens
Ann.Rev.Plant Physiol. 37, 509 - 538
- Mozzer,T.J. (1980)
Partial purification and characterization of the mRNA for α -amylase from barley aleurone layers
Plant Physiol. 65, 834 - 837
- Müller,G., Zimmermann,R. (1987)
Import of preproinsulin into the endoplasmic reticulum: structural basis for independence of SRP and docking protein
EMBO J. 6, 2099 - 2107
- Munro,S., Pelham,H.R.B. (1987)
A C-terminal signal prevents secretion of luminal ER proteins
Cell 48, 899 - 907
- Murashige,T., Skoog,F. (1962)
A rapid method for rapid growth and bioassays with tobacco tissue cultures
Physiol.Plant. 15, 473 - 497
- Nagy,J.L., Maliga,P. (1976)
Callus induction and plant regeneration from mesophyll protoplasts of *Nicotiana glauca*
Z.Pflanzenphysiol. 78, 433 - 435
- Negrotti,L., Shillito,R., Potrykus,L., Biasini,G., Sala,F. (1987)
Hybrid genes in the analysis of transformation conditions
Plant Mol.Biol. 8, 363 - 373
- Neuhaus,G., Neuhaus-Url,G., Schweiger,H.G., Chua,N.H., Broglie,R., Coruzzi,G., Chu,H. (1983)
Expression of higher plant genes in the green alga *Acetabularia mediterranea*
Eur.J.Cell Biol. Suppl. 2, 30
- Neuhaus,G., Neuhaus-Url,G., Gruss,P., Schweiger,H.G. (1984)
Enhancer-controlled expression of the simian virus 40 T-antigen in the green alga *Acetabularia*
EMBO J. 3, 2169 - 2172
- Neuhaus,G., Neuhaus-Url,G., de Groot,E.J., Schweiger,H.G. (1986)
High yield and stable transformation of the unicellular green alga *Acetabularia* by microinjection of SV 40 DNA and pSV 2neo
EMBO J. 5, 1437 - 1444
- Okker,R.J.H., Spaink,H., Hille,J., Van Brussel,T.A.N., Lugtenberg,B., Schilperoort,R.A. (1984)
Plant-inducible virulence promoter of the *Agrobacterium tumefaciens* Ti plasmid
Nature 312, 564 - 566
- Olden,K., Pratt,R.M., Yamada,K.M. (1978)
Role of carbohydrates in protein secretion and turnover: effects of tunicamycin on the major cell surface glycoprotein of chick embryo fibroblasts
Cell 13, 461 - 473
- Olden,K., Parent,J.B., White,S.L. (1982)
Carbohydrate moieties of glycoproteins.
A re-evaluation of their function
Biochim.Biophys.Acta 650, 209 - 232

- Oliver, D. (1985)
Protein secretion in *Escherichia coli*
Ann. Rev. Microbiol. 39, 615 - 648
- Otten, L., Schilperoort, R. (1978)
A rapid microscale method for detection of lysopine and nopaline dehydrogenase activities
Biochim. Biophys. Acta 521, 491 - 500
- Ow, D.W., Jacobs, J.W., Howell, S.H. (1981)
Functional regions of the cauliflower mosaic virus 35S RNA promoter determined by the use of the firefly luciferase gene as a reporter of promoter activity
Proc. Natl. Acad. Sci. USA 84, 4870 - 4874
- Owen, J.E., Schultz, D.W., Taylor, A., Smith, G.R. (1983)
Nucleotide sequence of the lysozyme gene of bacteriophage T4
J. Mol. Biol. 165, 229 - 248
- Palni, L.M.S., Burch, L., Horgan, R. (1988)
The effect of auxin concentration on cytokinin stability and metabolism
Planta 174, 231 - 234
- Parent, J.-G., Asselin, A. (1984)
Detection of pathogenesis-related (PR or b) and of other proteins in the intercellular fluid of hypersensitive plants infected with tobacco mosaic virus
Canad. J. Bot. 62, 564 - 569
- Paszkowski, J., Shillito, R.D., Saul, M., Mandak, V., Hohn, T., Hohn, B., Potrykus, I. (1984)
Direct gene transfer to plants
EMBO J. 3, 2717 - 2722
- Peralta, E.G., Hellmiss, R., H.J.M., Berger, W.H., Ream, W. (1986)
Overdrive, a T-DNA transmission enhancer of the *A. tumefaciens* tumor-inducing plasmid
in: *Mol. Genet. Plant-Microbe Interactions, Proc. Int. Symp. 3. Meeting*, S. 20 - 26
ed.: Verma, D.P.S., Brisson, N.
Nijhoff, Dordrecht, Niederlande
- Perry, L.J., Wetzel, R. (1984)
Disulfide bond engineered into T4 lysozyme
Science 226, 555 - 557
- Perry, L.J., Heyneker, H.L., Wetzel, R. (1985)
Non-toxic expression in *E. coli* of a plasmid-encoded gene for phage T4 lysozyme
Gene 38, 359 - 364
- Pilet, P.E., Jollès, J., Jollès, P. (1983)
Lysozyme in growing plant tissue cultured in vitro
Z. Pflanzenphysiol. 111, 91 - 94
- Pilet, P.E., Bernasconi, P. (1984)
Rubus hispidus callus cultured in vitro and endogenous lysozyme: growth processes
Plant Sci. Lett. 35, 147 - 151
- Porter, S.G., Yanofsky, M.F., Nester, E.W. (1981)
Molecular characterization of the *virD* operon from *Agrobacterium tumefaciens*
Nucl. Acids Res. 15, 7503 - 7518
- Prehn, L.J., Wiedmann, M., Rapoport, T.A., Zwieb, C. (1981)
Protein translocation across wheat germ microsomal membranes requires an SRP-like component
EMBO J. 6, 2093 - 2097
- Primke, M., Berger, S., Schweiger, H.G. (1978)
Protoplasts from *Acetabularia*: isolation and fusion
Cytobiol. 16, 375 - 380

- Prols, M. (1986)
Transient expression of chimeric genes in protoplasts – Gene transfer isolated DNA with different methods.
Thesis, Cologne University
- Redman, C.M., Sabatini, D.D. (1966)
Vectorial discharge of peptides released by puromycin from attached ribosomes
Proc.Natl.Acad.Sci.USA 56, 608 - 615
- Reiss, B., Sprengel, R., Will, H., Schaller, H. (1984)
A new and sensitive method for qualitative and quantitative assay of neomycinphosphotransferase in crude cell extracts
Gene 30, 211 - 216
- Remington, S.J., Matthews, B.W. (1978)
A general method to assess similarity of protein structures, with applications to T4 bacteriophage lysozyme
Proc.Natl.Acad.Sci.USA 75, 2180 - 2184
- Reth, M. (1981)
Characterization of individual antibodies and idiotopes of the NP^b idiotope.
Dissertation, Cologne University
- Rigby, P.W.J., Dieckmann, M., Rhodes, C., Berg, P. (1977)
Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA Polymerase I
J.Mol.Biol. 113, 237 - 251
- Riggleman, R.C., Fristensky, B., Hadwiger, L.A. (1985)
The disease resistance response in pea is associated with increased levels of specific mRNAs
Plant Mol.Biol. 4, 81 - 86
- Robinson, A., Austen, B.M. (1981)
The role of topogenic sequences in the movement of proteins through membranes
Biochem.J. 246, 249 - 261
- Robinson, A., Kaderbhai, M.A., Austen, B.M. (1987)
Identification of signal sequence binding proteins integrated into the rough endoplasmic reticulum membrane
Biochem.J. 242, 762 - 767
- Rogers, J.C., Millman, C. (1983)
Isolation and sequence analysis of a barley α -amylase cDNA clone
J.Biol.Chem. 258, 8169 - 8174
- Rothman, J.E. (1987)
Protein sorting by selective retention in the endoplasmic reticulum and Golgi stack
Cell 50, 521 - 522
- Rothstein, S.J., Lazarus, C.M., Smith, W.E., Baulcombe, D.C., Gatenby, A.A. (1984)
Secretion of a wheat α -amylase expressed in yeast
Nature 308, 662 - 665
- Rubin, R.A. (1986)
Genetic studies on the role of octopine T-DNA border regions in crown gall tumor formation
Mol.Gen.Genet. 202, 312 - 320
- Salomon, M.J., Strauss, F., Varshavsky, A. (1986)
A mammalian high mobility group protein recognizes any stretch of six AT base pairs in duplex DNA
Proc.Natl.Acad.Sci.USA 83, 1276 - 1280

- Sanchez-Serrano, J., Schmidt, R., Schell, J., Willmitzer, L. (1986)
Nucleotide sequence of proteinase inhibitor II encoding cDNA of potato (*Solanum tuberosum*) and its mode of expression
Mol.Gen.Genet. 203, 15 - 20
- Sanders, P.R., Winter, J.A., Barnason, A.R., Rogers, S.G., Fraley, R.T. (1987)
Comparison of cauliflower mosaic virus 35S and nopaline synthase promoters in transgenic plants
Nucl.Acids Res. 15, 1543 - 1558
- Schlumbaum, A., Mauch, F., Vögeli, U., Boller, T. (1986)
Plant chitinases are potent inhibitors of fungal growth
Nature 324, 365 - 367
- Schreier, P.H., Seftor, E.A., Schell, J., Bohnert, H.J. (1985)
The use of nuclear encoded sequences to direct the light regulated synthesis and transport of a foreign protein into plant chloroplasts
EMBO J. 4, 25 - 32
- Schröder, G., Waffenschmidt, S., Weller, E.W., Schröder, J. (1984)
The T-region of T1 plasmids codes for an enzyme synthesizing indole-3-acetic acid
Eur.J.Biochem. 138, 387 - 391
- Schweiger, H.-G., Neuhaus, G. (1987)
Induction of expression in and stable transformation of an algal cell by nuclear microinjection with naked DNA
in: *Plant Gene Research / Plant DNA Infectious Agents*
ed.: Hohn, T., Schell, J., S. 285 - 302
Springer Verlag Wien, New York
- Selsted, M.E., Martinez, R.J. (1980)
A simple and ultrasensitive enzymatic assay for the quantitative determination of lysozyme in the picogram range
Anal.Biochem. 109, 67 - 70
- Shah, D.M., Horsch, R.B., Klee, H.J., Kishore, G.M., Winter, J.A., Turner, N.E., Hiratake, C.M., Sanders, P.R., Gasser, C.S., Aykent, S., Siegel, N.R., Rogers, S.G., Fraley, R.T. (1986)
Engineering herbicide tolerance in transgenic plants
Science 233, 478 - 481
- Sheikholeslam, S.N., Weeks, D.P. (1987)
Acetoxysyringone promotes high efficiency transformation of *Arabidopsis thaliana* explants by *Agrobacterium tumefaciens*
Plant Mol.Biol. 8, 291 - 298
- Shillito, R.D., Saul, M.W., Paszkowski, J., Müller, M., Potrykus, I. (1985)
High efficiency direct gene transfer to plants
Biotechnol. 3, 1099 - 1103
- Shinshi, H., Mohnen, D., Meins, F. (1987)
Regulation of a plant pathogenesis-related enzyme: Inhibition of chitinase and chitinase mRNA accumulation in cultured tobacco tissues by auxin and cytokinin
Proc.Natl.Acad.Sci.USA 84, 89 - 93
- Sjöström, M., Wold, S., Wieslander, A., Rilfors, L. (1987)
Signal peptide amino acid sequences in *Escherichia coli* contain information related to final protein localization.
A multivariate data analysis
EMBO J. 6, 823 - 831
- Sosountzov, L., Sotta, B., Maldiney, R., Salbaugh, L., Miglino, E. (1986)
Immunoelectron-microscopy localization of abscisic acid with colloidal gold on Lowicryl-embedded tissues of *Chenopodium polyspermum* L.
Planta 168, 471 - 481
- Southern, E.M. (1975)
Detection of specific sequences among DNA fragments separated by gel electrophoresis
J.Mol.Biol. 98, 503 - 517

- Stachel,S.E., Nester,E.W., Zambryski,P. (1986a)
A plant cell factor induces *Agrobacterium tumefaciens* vir gene expression
Proc.Natl.Acad.Sci.USA 83, 379 - 383
- Stachel,S.E., Timmerman,B., Zambryski,P. (1986b)
Generation of single-stranded T-DNA molecules during the initial stages of
T-DNA transfer from *Agrobacterium* to plant cells
Nature 322, 706 - 712
- Stachel,S.E., Nester,E.W. (1986c)
The genetic and transcriptional organization of the vir region of the
A6 Ti plasmid of *Agrobacterium tumefaciens*
EMBO J. 5, 1445 - 1454
- Stachel,S.E., Zambryski,P. (1986d)
Agrobacterium tumefaciens and the susceptible plant cell: a novel
adaptation of extracellular recognition and DNA conjugation
Cell 47, 1155 - 157
- Stachel,E.W., Zambryski,P. (1986e)
virA and virG control the plant-induced activation of the T-DNA transfer
process of *Agrobacterium tumefaciens*
Cell 46, 325 - 333
- Staehelin,L.A., Chapman,R.L. (1987)
Secretion and membrane recycling in plant cells: novel intermediary
structures visualized in ultrarapidly frozen sycamore and carrot
suspension-culture cells
Planta 171, 43 - 57
- Stalker,D.M., Hiatt,W.R., Comai,L. (1985)
A single amino acid substitution in the enzyme 5-enolpyruvylshikimate-3-
phosphate synthase
J.Biol.Chem. 260, 4724 - 4728
- Steinhilber,H.C., Amrhein,N. (1980)
The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic
acid-3-phosphate synthase
Biochem.Biophys.Res.Comm. 94, 1207 - 1212
- Stieger, M. (1987)
Experiments on integration and expression of chimeric immunoglobulins in plants.
Dissertation, Cologne University
- Szczesna-Skorupa,E., Browne,N., Mead,D., Kemper,B. (1988)
Positive charges at the NH₂ terminus convert the membrane-anchor signal
peptide of cytochrome P-450 to a secretory signal peptide
Proc.Natl.Acad.Sci. USA 85, 738 - 742
- Tijssen,P. (1985)
Practice and theory of enzyme immunoassays, pp. 79 - 84
Elsevier, Amsterdam, New York, Oxford
- Thompson,C.J., Rao Movva,N., Tizard,R., Cramer,E., Davies,J.E.,
Lanwerreys,M., Botterman,J. (1987)
Characterization of the herbicide-resistance gene bar from *Streptomyces*
hygroscopicus
EMBO J. 6, 2519 - 2523
- Towbin,H., Staehelin,T., Gordon,J. (1979)
Electrophoretic transfer of proteins from polyacrylamide gels to
nitrocellulose sheets: procedure and some applications
Proc.Natl.Acad.Sci.USA 76, 4350 - 4354
- Trugita,A., Inouye,M., Terzaghi,E., Streisinger,G. (1968)
Purification of bacteriophage T4 lysozyme
J.Biol.Chem. 243, 392 - 397
- Trugita,A. (1971)
Phage lysozyme and other lytic enzymes
in: Boyer,P.D. (ed.): *The Enzymes*
3. Auflage, Band V, S. 344 - 411
Academic Press, New York

- Uchimiya,H., Hirochika,H., Hashimoto,H., Hara,A., Masuda,T., Kasumimoto,T., Harada,H., Ikeda,J.-E., Yoshika,M. (1986)
Co-expression and inheritance of foreign genes in transformants obtained by direct DNA transformation to tobacco protoplasts
Mol.Gen.Genet. 205, 1 - 8
- Uzami,S., Morikawa,S., Takebe,I., Machida,Y. (1987)
Absence in monocotyledonous plants of the diffusible plant factors inducing T-DNA circularization and vir gene expression in *Agrobacterium*
Mol.Gen.Genet. 209, 221 - 226
- Vaeck,M., Reynaerts,A., H6fte,H., Jansens,S., De Beuckeleer,M., Dean,C., Zabeau,M., Van Montagu,M., Leemans,J. (1987)
Transgenic plants protected from insect attack
Nature 328, 33 - 37
- Van Haeren,M.J.J., Pronk,J.T., Schilperoort,R.A., Hooykaas,P.J.J. (1987)
Functional analysis of the *Agrobacterium tumefaciens* octopine Ti-plasmid left and right T-region border fragments
Plant Mol.Biol. 8, 95 - 104
- Van Haute,E., Joos,H., Maes,M., Warren,G., Van Montagu,M., Schell,J. (1983)
Intergenic transfer and exchange recombination of restriction fragments cloned in pBR 322: a novel strategy for the reversed genetics of Ti-plasmids of *A.tumefaciens*
EMBO J. 2, 411 - 418
- Van Loon,L.C. (1985)
Pathogenesis-related proteins
Plant Mol.Biol. 4, 111 - 116
- Van Loon,L.C., Gerritsen,Y.A.M., Ritter,C.E. (1987)
Identification, purification and characterization of pathogenesis-related proteins from virus-infected Samsun NN tobacco leaves
Plant Mol.Biol. 9, 593 - 609
- Van Onckelen,H., Prinsen,E., Inzé,D., Rüdelsheim,P., Van Lijsebettens,M., Follin,A., Schell,J., Van Montagu,M., De Greef,J. (1986)
Agrobacterium T-DNA gene 1 codes for tryptophan-2-monooxygenase activity in tobacco crown gall cells
FEBS Lett. 198, 357 - 360
- Velten,J., Velten,L., Hain,R., Schell,J. (1984)
Isolation of a dual plant promoter fragment from the Ti plasmid of *Agrobacterium tumefaciens*
EMBO J. 3, 2723 - 2730
- Velten,J., Schell,J. (1986)
Selection-expression plasmid vectors for use in genetic transformation of higher plants
Nucl.Acids Res. 13, 6981 - 6998
- Veluthambi,K., Jayaswal,R.K., Gelvin,S.B. (1987)
The virulence genes A, G and D mediate the double-stranded border cleavage of the T-DNA from the *Agrobacterium tumefaciens* Ti plasmid
Proc.Natl.Acad.Sci.USA 84, 1881 - 1885
- Vieira,J., Messing,J. (1982)
The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers
Gene 19, 259 - 268
- von Heijne,G. (1983)
Patterns of amino acids near signal sequence cleavage sites
Eur.J.Biochem. 113, 17 - 21
- von Heijne,G. (1984)
Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in procaryotic and eucaryotic cells
EMBO J. 3, 2315 - 2318
- von Heijne,G. (1985a)
Signal sequences - The limits of variation
J.Mol.Biol. 184, 99 - 105
- von Heijne,G. (1985b)
Ribosome - SRP - signal sequence interactions
FEBS Lett. 190, 1 - 5

- von Heijne, G. (1986a)
Net N-C - charge imbalance may be important for signal sequence function in bacteria
J. Mol. Biol. 192, 287 - 290
- von Heijne, G. (1986b)
A new method for predicting signal sequence cleavage sites
Nucl. Acids Res. 14, 4683 - 4690
- Wahl, G.M., Stern, M., Stark, G. (1979)
Efficient transfer of large DNA fragments from agarose gels to diazobenzylmethyl-paper and rapid hybridization by using dextran sulfate
Proc. Natl. Acad. Sci. USA 76, 3683 - 3687
- Walter, P., Blobel, G. (1981)
Translocation of proteins across the endoplasmic reticulum.
III. Signal recognition particle (SRP) causes signal-sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes
J. Cell Biol. 91, 557 - 561
- Walter, P., Gilmore, R., Blobel, G. (1984)
Protein translocation across the endoplasmic reticulum
Cell 38, 5 - 8
- Wang, K., Stachel, S.E., Timmerman, B., Van Montagu, M., Zambryski, P.C. (1987a)
Site-specific nick in the T-DNA border sequence as a result of *Agrobacterium* vir gene expression
Science 235, 587 - 591
- Wang, K., Genetello, C., Van Montagu, M., Zambryski, P. (1987b)
Sequence context of the T-DNA border repeat elements determines the relative activity during T-DNA transfer to plant cells
Mol. Gen. Genet. 210, 338 - 346
- Walter, P. (1987)
Signal recognition: Two receptors act sequentially
Nature 328, 763 - 764
- Weller, R.W., Schröder, J. (1987)
Hormone genes and crown gall disease
Trends Biotechnol. Sci. 12, 271 - 275
- Wetzel, R., Perry, L.J., Baase, W.A., Becktel, W.J. (1988)
Disulfide bonds and thermal stability in T4 lysozyme
Proc. Natl. Acad. Sci. USA 85, 401 - 405
- Wickens, M., Stephenson, P. (1984)
Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3' end formation
Science 226, 1045 - 1051
- Wiedmann, M., Kurzchalia, T.V., Bielka, H., Rapoport, T.A. (1987a)
Direct probing of the interaction between the signal sequence of nascent preprolactin and the signal recognition particle by specific cross-linking
J. Cell Biol. 104, 201 - 208
- Wiedmann, M., Kurzchalia, T.V., Hartmann, E., Rapoport, T.A. (1987b)
A signal sequence receptor in the endoplasmic reticulum membrane
Nature 328, 830 - 833
- Winter, J.A., Wright, R.L., Gurley, W.B. (1984)
Map location of five transcripts homologous to T_R-DNA in tobacco and sunflower crown gall tumors
Nucl. Acids Res. 12, 2391 - 2406
- Wood, C.R., Boss, M.A., Kenten, J.H., Calvert, J.E., Roberts, N.A., Emtage, J.S. (1985)
The synthesis and in vivo assembly of functional antibodies in yeast
Nature 314, 446 - 449

- Wu,H.C., Tai,P.C. (1986)
Current topics in Microbiology and Immunology 125 - Protein Secretion and Export in Bacteria
Springer Verlag Berlin, Heidelberg, New York, Tokio
- Yamamoto,A., Iwahashi,M., Yanofsky,M.F., Nester,E.W., Takebe,I.T., Machida,Y. (1987)
The promoter proximal region in the virD locus of *Agrobacterium tumefaciens* is necessary for the plant-inducible circularization of T-DNA
Mol.Gen.Genet. 206, 174 - 177
- Yanish-Perron,C., Viera,J., Messing,J. (1985)
Improved M13 cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC 19 vectors
Gene. 33, 103 - 119
- Yanofsky,M., Lowe,B., Montoya,A., Rubin,R., Krul,W., Gordon,M., Nester,E. (1985)
Molecular and genetic analysis of factors controlling host range in *Agrobacterium tumefaciens*
Mol.Gen.Genet. 201, 237 - 246
- Zambryski,P., Joos,H., Genetello,C., Van Montagu,M., Schell,J. (1983)
Ti-plasmid vector for introduction of DNA into plant cells without altering their normal regeneration capacity
EMBO J. 2, 2143 - 2150
- Zenke,M., Grundstrom,T., Matther,H., Wintzerith,M., Schatz,C., Wildemann,A., Chambon,P. (1986)
Multiple sequence motifs are involved in SV 40 enhancer function
EMBO J. 3, 387 - 397
- Zimmermann,R., Meyer,D.I. (1986)
1986: A year of new insights into how proteins cross membranes
Trends Biotechnol.Sci. 11, 512 - 515

THIS PAGE BLANK (USPTO)

Posttranslational Association of Immunoglobulin Heavy Chain Binding Protein with Nascent Heavy Chains in Nonsecreting and Secreting Hybridomas

David G. Bole, Linda M. Hendershot, and John F. Kearney

Cellular Immunobiology Unit of the Tumor Institute, Department of Microbiology, and the Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, Alabama 35294

Abstract. A rat monoclonal antibody specific for immunoglobulin (Ig) heavy chain binding protein (BiP) has allowed the examination of the association of BiP with assembling Ig precursors in mouse B lymphocyte-derived cell lines. The anti-BiP monoclonal antibody immunoprecipitates BiP along with noncovalently associated Ig heavy chains. BiP is a component of the endoplasmic reticulum and binds free intracellular heavy chains in nonsecreting pre-B (μ^+ , L^-) cell lines or incompletely assembled Ig precursors in (H^+ , L^+) secreting hybridomas and myelomas. In the absence of light chain synthesis, heavy chains remain associated with BiP and are not secreted. The associa-

tion of BiP with assembling Ig molecules in secreting hybridomas is transient and is restricted to the incompletely assembled molecules which are found in the endoplasmic reticulum. BiP loses affinity and disassociates with Ig molecules when polymerization with light chain is complete. We propose that the association of BiP with Ig heavy chain precursors is a novel posttranslational processing event occurring in the endoplasmic reticulum. The Ig heavy chains associated with BiP are not efficiently transported from the endoplasmic reticulum to the Golgi apparatus. Therefore, BiP may prevent the premature escape and eventual secretion of incompletely assembled Ig molecules.

THE process by which B lymphocytes synthesize immunoglobulin (Ig)¹ involve the assembly of heavy and light chains, glycosylation, posttranslational modification of oligosaccharides, and transport between intracellular compartments before eventual secretion (41). Ig appears to be transported along the same intracellular pathway as other secretory glycoproteins.

Comparative examination of the intracellular transport of many secretory proteins indicates that transport from the rough endoplasmic reticulum (RER) to the Golgi apparatus is selective and can be the rate-limiting step in the secretory process (11, 24). This selectivity and the demonstration of different rates of transport between the RER and Golgi apparatus for different secretory proteins within a cell line has led to the hypothesis that secretory proteins express specific transport signals within their structure (11, 24) and that this transport may be receptor mediated (12, 24, 40).

Expression and selective intracellular transport of Ig in B-lymphocyte lineage cells is developmentally regulated. Pre-B cell lymphocytes differentiate from bone marrow multipotential stem cells and are first identified by the presence of cytoplasmic μ -heavy chains in the absence of light chain synthesis (8, 27). The μ -chains of bone marrow derived pre-

B cells are of the membrane type but are not expressed on the cell surface (42). The onset of light chain synthesis coincides with and appears to be necessary for the expression of surface and secretory Ig (26).

Mouse pre-B-derived hybridomas and Abelson-virus transformed pre-B cell lines synthesize membrane and secretory μ -chains which remain intracellular and are not secreted or expressed on the cell surface (17, 32, 37). The analysis of intracellular μ -chains from pre-B derived cell lines as well as variant nonsecreting myelomas has identified a protein of 78,000 daltons, Ig heavy chain binding protein (BiP), that coimmunoprecipitates with Ig heavy chains (13, 28). Because BiP was readily observed in cell lines representative of the pre-B cell stage of development, others have speculated that BiP functions in the regulation of Ig synthesis or prevents heavy chain insolubility in pre-B cells (43).

The results of the presently described studies demonstrate that BiP binds free heavy chains as they occur in pre-B derived cell lines, and incompletely assembled Ig molecules in secreting cell lines. Our interpretation of these observations are that BiP plays a role in the posttranslational processing of nascent Ig heavy chains in the endoplasmic reticulum. Heavy chains associated with BiP are not efficiently transported from the RER to the Golgi apparatus. This suggests that the posttranslational interaction of BiP with Ig heavy chains interferes with the expression of transport signals inherent in the Ig molecule.

1. Abbreviations used in this paper: BiP, immunoglobulin heavy chain binding protein; Ig, immunoglobulin; RER, rough endoplasmic reticulum; SER, smooth endoplasmic reticulum.

We
blec
plas
plet
asse
secre

Mc

Rea

Affin
obtai
[³⁵S],
Suga
Dolic
St. L
San I
GIBC
sourc

Cel.

Cell l
with
 μ M),
cell-d
P3-X
ratio
of Co
of P3
MM
speci
whic

Pre.

The
immu
lysed
200 μ
S. au
buffer
S. au
eluate
resus
third
noge
in ph
popli
mas
reacti
secre
assay
label

Rac

Before
5 x 1
serum
prein
cultu
pulse
adde
state
by ce
lysis
and
short
and
centr
direct
indir
prote

Bole,



THIS PAGE BLANK (USPTO)

We propose that association of B μ with incompletely assembled Ig molecules prevents their transport from the endoplasmic reticulum. Once assembly with light chain is complete, disassociation of BiP occurs, allowing transport of the assembled molecule to the Golgi apparatus and eventual secretion.

Materials and Methods

Reagents

Affinity-purified isotype specific antibodies used for immunoprecipitation were obtained from Southern Biotechnology Associates Inc., Birmingham, AL. [³⁵S]Methionine was purchased from Amersham Corp., Arlington Heights, IL. Sugar nucleotides were obtained from New England Nuclear, Boston, MA. Dolicol phosphate and protein A were purchased from Sigma Chemical Co., St. Louis, MO, and tunicamycin was obtained from Calbiochem-Behring Corp., San Diego, CA. Tissue culture media and fetal calf serum were purchased from GIBCO, Grand Island, NY. All other reagents were obtained from different sources and were of analytical grade.

Cell Culture

Cell lines were maintained in stationary culture in RPMI 1640 supplemented with L-glutamine (2 mM), penicillin-streptomycin, β -mercaptoethanol (50 μ M), fungizone 1:1,000, and 15% fetal calf serum at 37°C in 10% CO₂. Pre-B cell-derived hybridoma 15-58 was obtained by fusion of 19-d fetal liver with P3-X63-Ag8.653, as previously described (8). Ag8(8) was isolated in collaboration with A. Radbruch and K. Rajewsky (Department of Genetics, University of Cologne, FRG), by fluorescent-activated cell sorting as a spontaneous variant of P3X63-Ag8, and produces only cytoplasmic γ_1 heavy chains. RD3-2 and MM60 are conventional Ig-secreting mouse hybridomas with anti-idiotypic specificity (33, 38). 11-11 is a rat lymph node x Ag8.653 heterohybridoma which secretes rat IgG with anti-mouse IgD specificity.

Preparation of Monoclonal Anti-BiP

The immunogen containing mouse BiP: μ -chain complexes was obtained by immune precipitation of 5×10^7 15-58 pre-B hybridoma cells. The cells were lysed in 15 ml of 1% Nonidet P-40 lysis buffer and immunoprecipitated with 200 μ g of goat anti-mouse μ -chain antibody and *Staphylococcus aureus*. The *S. aureus* pellet was washed three times with 0.75 M guanidinium hydrochloride buffered in 10 mM Tris, pH 7.5. The immune complexes were eluted from the *S. aureus* with 2.5 M buffered guanidinium hydrochloride. Proteins in the eluate were precipitated with 9 volumes of ethanol at -20°C for 1 h and resuspended in PBS for injection. A Fisher rat was immunized five times, every third day, in one rear footpad with this antigen. The first injection of immunogen was emulsified in complete Freund's adjuvant the others administered in phosphate-buffered saline (PBS). 1 d after the last injection the draining popliteal lymph node was fused with mouse myeloma Ag8.653 (16). Hybridomas were initially screened by an enzyme-linked immunosorbent assay for reactivity against both BiP: μ -chain complex and purified IgM. The hybridomas secreting antibody specific only for BiP by the enzyme-linked immunosorbent assay were further screened for their ability to immunoprecipitate BiP from a labeled cell lysate as detected by SDS PAGE.

Radiolabeling of Cell Lines

Before labeling with [³⁵S]methionine, cells were preincubated at a density of 1×10^6 /ml in methionine-free RPMI 1640 containing 15% dialyzed fetal calf serum for 30 min. When tunicamycin was used, it was included in the preincubation media at a final concentration of 2.5 μ g/ml from stock. Cell cultures were then labeled with 10 μ Ci of [³⁵S]methionine 1,400 Ci/mmol. In pulse-chase experiments, either complete medium or cold methionine was added at the time of chase to a concentration of 30 μ g/ml. Unless otherwise stated, cells were labeled 10-15 min before lysis or chase. Cells were harvested by centrifugation at 1,000 g for 5 min and lysed for 10 min at 0°C in 1 ml of lysis buffer containing 1% Nonidet P-40, 0.15 M NaCl, 50 mM Tris, pH 7.5, and 2 mM phenylmethylsulfonyl fluoride. In pulse-chase experiments with short chase periods, 2 \times lysis buffer was added directly to the cell culture. Nuclei and cellular debris were then removed by centrifugation in a Brinkmann centrifuge for 10 min. Supernatant was removed and immunoprecipitated by direct binding with monoclonal anti-BiP covalently linked to Sepharose 4B or indirectly with 50 μ l of anti-BiP culture supernate followed by addition of protein A-Sepharose. Ig heavy chains were immunoprecipitated with 10 μ g of

affinity-purified isotype specific goat anti-mouse antibody and protein A-Sepharose. Protein A or anti-BiP antibody, 1.0 mg per ml Sepharose, was coupled to CNBr-activated Sepharose 4B (31). Lysates were allowed to incubate with precipitating antibody for 1 h at 4°C with gentle mixing. Sepharose-bound immune complexes were washed three times with lysis buffer containing 1% Nonidet P-40, 0.5% deoxycholate, 0.1% SDS, 0.5 M NaCl, 50 mM Tris, pH 7.5, and once with PBS to remove detergents and salt. Precipitated proteins were then boiled in SDS sample buffer and subjected to SDS PAGE (21). Before autoradiography gels were fixed and then impregnated with ENHANCE (New England Nuclear) as directed by the manufacturer. Kodak X-omat AR X-ray film was used to visualize labeled proteins.

Subcellular Fractionation

Subcellular fractionation was performed by a modification of an unpublished procedure (Gary Sahagian, Department of Physiology, Tufts University, Boston, MA, personal communication). 5×10^7 RD3-2 or Ag8(8) hybridoma cells were harvested by centrifugation at 1,000 g for 5 min at 4°C, and first washed twice with PBS and then twice with 0.25 M sucrose in 5 mM Hepes buffered to pH 6.8. The cells were then resuspended in 2 ml of isotonic sucrose and transferred to a nitrogen cavitation bomb (Kontes Co., Vineland, NJ). The cells were pressurized to 40 psi for 30 min at 0°C. After decompression the cells were dispersed with 10 gentle strokes in a tight-fitting dounce homogenizer. Nuclei and cellular debris were removed by centrifugation at 800 g for 10 min. The postnuclear supernate was then layered on top of a discontinuous sucrose gradient containing 1 ml/2.0 M, 3.4 ml/1.3 M, 3.4 ml/1.0 M, and 2.75 ml/0.6 M sucrose in 5 mM Hepes (pH 6.8) in a Beckman 14 \times 89 mm Ultra-Clear tube. After 2 h of centrifugation at 40,000 rpm in a Beckman SW 41 rotor, ~19 fractions containing 10 drops each were collected from the bottom of the tubes. Each fraction was then assayed for endoplasmic reticulum and Golgi marker enzymes.

Marker Enzyme Assays

Mannosyltransferase was measured by methods as previously described (39). Membrane fractions were assayed in a total volume of 80 μ l containing 5×10^4 cpm GDP-mannose [mannose-3,4-³H(N)] 40-60 Ci/mmol, 0.05 mg dolicol phosphate, 1.25 mM MgCl₂, 0.015% Triton X-100, 20 mM Tris, pH 7.5, and 0.15 M NaCl. The assay mix was sonicated before addition of suitably diluted membrane fractions. Assays were incubated for 10 min at room temperature. Reactions were terminated by addition of 1.5 ml of chloroform/methanol (2:1), followed by 0.3 ml of saline. The tubes were vortexed, and chilled on ice for 10 min. Each tube was then spun for 5 min at 2,000 g and 0.6 ml of the chloroform was removed, dried in scintillation vials, and counted in scintillation cocktail.

Galactosyltransferase was measured by a modification of a procedure previously described (7). Membrane fractions were incubated in a final volume of 40 μ l containing 5×10^4 cpm UDP-Galactose [Galactose-¹⁴C(U)] 337.0 mCi/mmol, 0.28 mg ovalbumin (Sigma Chemical Co., grade V), 50 mM Tris, pH 7.5, 20 mM MnCl₂, 0.5% Triton X-100. Reactions were incubated at 37°C for 60 min. 30 μ l of each assay was then spotted on Whatman 3-MM filter paper and dried at room temperature. The filter paper was then washed twice in 10% trichloroacetic acid, 2% phosphotungstic acid, and 1% pyrophosphate at 0°C for 15 min. The filter paper was then washed in methanol, dried, and counted.

Identification of Immunoglobulin and BiP in Subcellular Fractions

BiP and Ig were identified in subcellular fractions by immunoprecipitation with monoclonal anti-BiP or with goat anti-heavy chain specific antibody. 1×10^6 [³⁵S]methionine-labeled cells were fractionated by the above procedure. Before immunoprecipitation each fraction was diluted with an equal volume of 2 \times lysis buffer to solubilize membranes. Each fraction was then divided for precipitation with anti-BiP or anti-Ig heavy chain.

Results

Immunoprecipitation of BiP and Immunoglobulin Heavy Chains from a Nonsecreting (γ_1^+ , κ^-) and a Secreting (γ_1^+ , κ^+) B-Cell Hybridoma

To compare the association of BiP with free and assembling heavy chains we have selected a nonsecreting variant of the myeloma P3X63-Ag8, designated Ag8(8) (γ_1^+ , κ^-), and a secreting hybridoma RD3-2 (γ_1^+ , κ^+). Like pre-B cell-derived

THIS PAGE BLANK (USFTQ)

hybridomas, Ag8(8) produces only heavy chains which remain intracellular. Cells were labeled with [³⁵S]methionine for 15 min and lysed immediately with detergent or chased for 60 min in excess cold methionine before lysis. The lysates were divided and immunoprecipitated separately with monoclonal anti-BiP or anti- γ_1 heavy chain-specific antibody. When Ig-secreting RD3-2 cells were immunoprecipitated with anti- γ_1 antibody and the precipitated proteins separated by SDS PAGE under reducing conditions, heavy chains and light chains were detected (Fig. 1A, lane 1). When cell lysates from the 15-min pulse were immunoprecipitated with anti-BiP antibody, BiP with *M_r* 78,000 and a band migrating with the same mobility of γ_1 -heavy chain were detected (Fig. 1A, lane 2). No detectable band co-migrating with light chains is seen in the anti-BiP immunoprecipitate. After a 60-min chase, immunoprecipitation with anti-BiP yielded the same *M_r* 78,000 band, however, little protein with the molecular weight of γ_1 heavy chain is observed (Fig. 1A, lane 4). After the 60-min chase, labeled intracellular heavy and light chains are still present and are precipitated with anti- γ_1 antibody (Fig. 1, lane 3). These results suggest that heavy chains in RD3-2 associate transiently with BiP after their synthesis, and these heavy chains that are associated with BiP have not acquired light chains. When the same type of experiment was conducted with nonsecreting Ag8(8), BiP and γ_1 heavy chains were precipitated as a complex regardless of which precipitating antibody was used (Fig. 1B). In contrast to the results obtained with RD3-2, the BiP heavy chain complex is stable and does not dissociate after a 60-min chase (Fig. 1B, lanes 3 and 4).

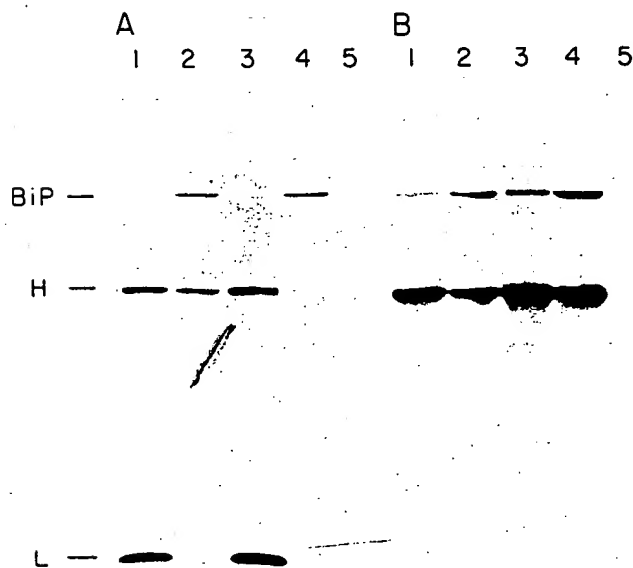


Figure 1. Association of BiP with γ_1 heavy chains in secreting RD3-2 and nonsecreting Ag8(8) hybridomas. Secreting RD3-2 cells (A) or nonsecreting Ag8(8) cells (B) were pulsed 15 min with [³⁵S]methionine, lysed in detergent, and immunoprecipitated with goat anti- γ_1 heavy chain antibody and protein A-Sepharose (lanes 1) or with monoclonal rat anti-BiP antibody coupled to Sepharose (lanes 2). Cells pulsed 15 min and chased 1 h in excess cold methionine before lysis were immunoprecipitated with goat anti- γ_1 heavy chain antibody (lanes 3) and anti-BiP Sepharose (lanes 4). Lanes 5 show immunoprecipitation of cell lysates after a 15-min label with a monoclonal rat anti-mouse IgD (11-11) as control. Immunoprecipitated proteins were reduced and resolved on 10% SDS PAGE.

A similar pattern of association between heavy chains and BiP was demonstrated in cells producing μ -chains. A nonsecreting pre-B-derived hybridoma, 15-58 which produces μ -chains in the absence of light chain synthesis and a conventional μ, λ secreting hybridoma MM60 were immunoprecipitated with anti- μ heavy chain-specific and anti-BiP antibody. The μ -chains from each of these cell lines migrate as two primary bands reflecting an apparent glycosylation heterogeneity that is observed among IgM oligosaccharides (1). After treatment with endoglycosidase H or when cells are labeled in the presence of tunicamycin a single band representing nonglycosylated μ -chain is resolved (Bole, D. G., L. M. Hendershot, J. F. Kearney, unpublished observations). After 15 min of labeling two bands with apparent mobility characteristic of μ -chains were immunoprecipitated from 15-58 with either anti-BiP or anti- μ chain antibodies (Fig. 2A, lanes 2 and 3). However, after a 1-h chase period, three bands are immunoprecipitated with both anti- μ and anti-BiP antibodies (Fig. 2A, lanes 4 and 5); the top band likely corresponds to BiP and the two bands below to μ -chains which now migrate with increased mobility. An increase in the mobility of the μ -chains could result from exoglycosidase trimming of the five oligosaccharide units of the μ -chains (15, 41). In contrast to 15-58, when lysates from MM60 were immunoprecipitated with anti-BiP antibody, very little co-precipitation of μ -chain is detected in the pulse and the chase (Fig. 2B, lanes 3 and 5). MM60 μ -chains immunoprecipitated with anti- μ antibody show an increase in mobility after the 60-min chase, similar to that observed in 15-58. Thus a prolonged association occurs between BiP and the μ -chains in nonsecreting 15-58 that is not observed in secreting MM60; similar to the results obtained with Ag8(8) and RD3-2.

Association of BiP with Assembling Immunoglobulin Precursors in RD3-2

Fig. 1A demonstrated that nascent heavy chains in RD3-2 that were co-precipitated with BiP contained few if any light

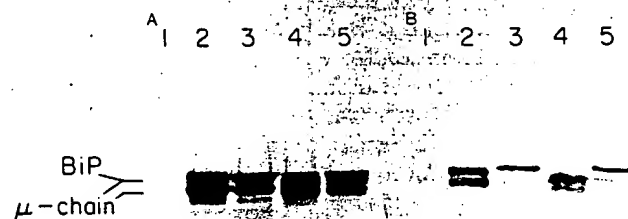


Figure 2. Association of BiP with μ -heavy chains in 15-58, a nonsecreting pre-B hybridoma and MM60, a conventional secreting hybridoma. Nonsecreting 15-58 cells (A) and secreting MM60 cells (B) were labeled 15 min with [³⁵S]methionine and immediately lysed in detergent and immunoprecipitated with goat anti- μ heavy chain antibody and protein A-Sepharose (lanes 2) or monoclonal anti-BiP antibody and protein A-Sepharose (lanes 3). Cells labeled for 15 min and then chased for 1 h were immunoprecipitated with goat anti- μ antibody and protein A-Sepharose (lanes 4) or monoclonal anti-BiP antibody and protein A (lanes 5). Cells labeled for 15 min and immunoprecipitated with protein A-Sepharose alone are shown in lanes 1. Immunoprecipitated proteins were reduced and resolved on 8% SDS PAGE.

THIS PAGE BLANK (USPTO)

chains. This observation suggested the association of BiP with heavy chains in this secreting hybridoma was a transient posttranslational event occurring before light chain addition. To further investigate this association, we examined the assembly of Ig precursors in RD3-2 by immunoprecipitation with anti- γ_1 heavy chain-specific and anti-BiP antibody. Pulse-chase experiments were performed and immunoprecipitated proteins resolved on nonreducing gels. When heavy chains were immunoprecipitated with anti- γ_1 antibody free heavy chains (H), heavy chain dimers (H₂), dimers with one light chain (H₂L), and complete (H₂L₂) molecules were observed (Fig. 3A). A small amount of BiP was also detected when lysates were immunoprecipitated with anti-heavy chain antibody. The preferred sequence of assembly of mouse IgG₁ requires heavy chains to polymerize into heavy chain dimers before light chain addition occurs (5). With increasing chase time incompletely assembled Ig precursors polymerized with light chains into the mature completely assembled H₂L₂ molecules. When immunoprecipitation was performed with anti-BiP antibody, free heavy chains, heavy chain dimers, and dimers with one light chain were co-precipitated with BiP. However, no completely assembled H₂L₂ molecules were associated with BiP (Fig. 3B). Thus when Ig assembly is complete affinity of BiP for RD3-2 heavy chains is lost.

The Effect of Tunicamycin on Immunoglobulin Assembly and Association with BiP

Before polymerization with light chain, Ig heavy chains are glycosylated co-translationally (3). Inhibiting the glycosylation of Ig heavy chains with tunicamycin could affect the folding of the nascent molecule within the endoplasmic reticulum. We examined the assembly of heavy and light chains in J558 an α, λ myeloma and compared the association of BiP with J558 heavy chains in cells which had been cultured in the presence and absence of tunicamycin. Intracellular transport and secretion of IgA in myelomas is sensitive to but not completely blocked by tunicamycin treatment (14).

J558 cells were pulse labeled and chased for increasing periods of time in the presence or absence of tunicamycin before lysis. Immunoprecipitation of Ig from these cells with anti- α heavy chain-specific antibody precipitated heavy

chains with associated light chains (Fig. 4, A and C). At early time points after pulse labeling, very few light chains co-precipitated with heavy chains in lysates from the tunicamycin-treated cultures (Fig. 4C, lanes 1 and 2). In contrast, light chains were rapidly associated with heavy chains in control cultures (Fig. 4A, lanes 1 and 2). These results suggest that subunit assembly occurs less efficiently when glycosylation is inhibited or blocked. A lack of heavy chain oligosaccharide units has been shown to inhibit assembly of Ig in vivo (10) and in vitro (18). If displacement of BiP is a consequence of light chain assembly, it would be anticipated that BiP association would be enhanced in the presence of tunicamycin.

Fig. 4, B and D, shows that BiP and associated heavy chains are precipitated with anti-BiP antibody. The amount of heavy chains co-precipitating with anti-BiP antibody is much greater when tunicamycin is included in the culture (compare Fig. 4, B and D). The heavy chains precipitated with BiP from the control lysate contain multiple bands which may reflect glycosylation heterogeneity of the α -heavy chains. The relative enhancement of BiP heavy chain complex observed in the tunicamycin-treated cells was estimated by densitometer tracings of heavy chains co-precipitated with BiP (Fig. 4B, lane 1 [sum of three species] and Fig. 4D, lane 1 [one species]). Based on band intensity of the autoradiographs, densitometer analysis revealed that four times as many heavy chains coprecipitated with anti-BiP antibody from lysates of tunicamycin-treated cells compared to the control cell cultures without tunicamycin. The relative quantity of heavy chain co-precipitating with BiP was standardized for the densitometer value of BiP precipitated from each fraction. Thus, in J558, prevention of glycosylation by tunicamycin appears to delay the assembly of heavy and light chains and retards the displacement of BiP from nonglycosylated heavy chains. These results suggest strongly that there exists a direct correlation between assembly of Ig and BiP displacement.

Autoimmune Nature of Anti-BiP Antibody

The results of the above-described studies suggest that BiP is associated with heavy chains after their synthesis and before addition of light chain is complete. This requires that BiP, and partially and completely assembled functional Ig mole-

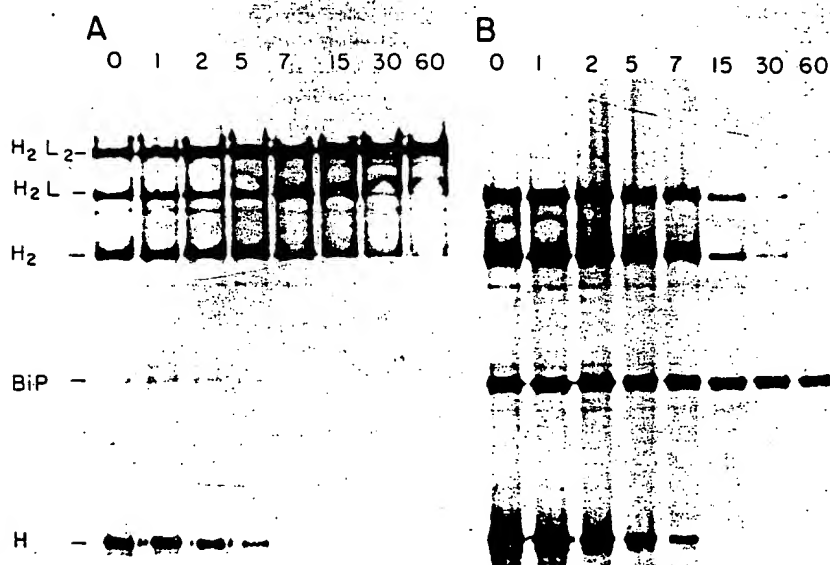


Figure 3. Association of BiP with assembling immunoglobulin precursors in RD3-2. RD3-2 cells were labeled for 10 min with [³⁵S]methionine and then chased with excess cold methionine for 0, 1, 2, 5, 7, 10, 20, and 60 min, at which time point an aliquot of the cell culture was lysed in detergent. Cell lysates from each time point were split and immunoprecipitated separately with either goat anti- γ_1 heavy chain antibody and protein A-Sepharose (A) or with monoclonal anti-BiP coupled to Sepharose (B). Assembling protein precursors were run on 8% SDS PAGE under nonreducing conditions. The fluoroenhanced gel represented in B was exposed to film approximately twice as long as the gel in A.

THIS PAGE BLANK (USPTO)

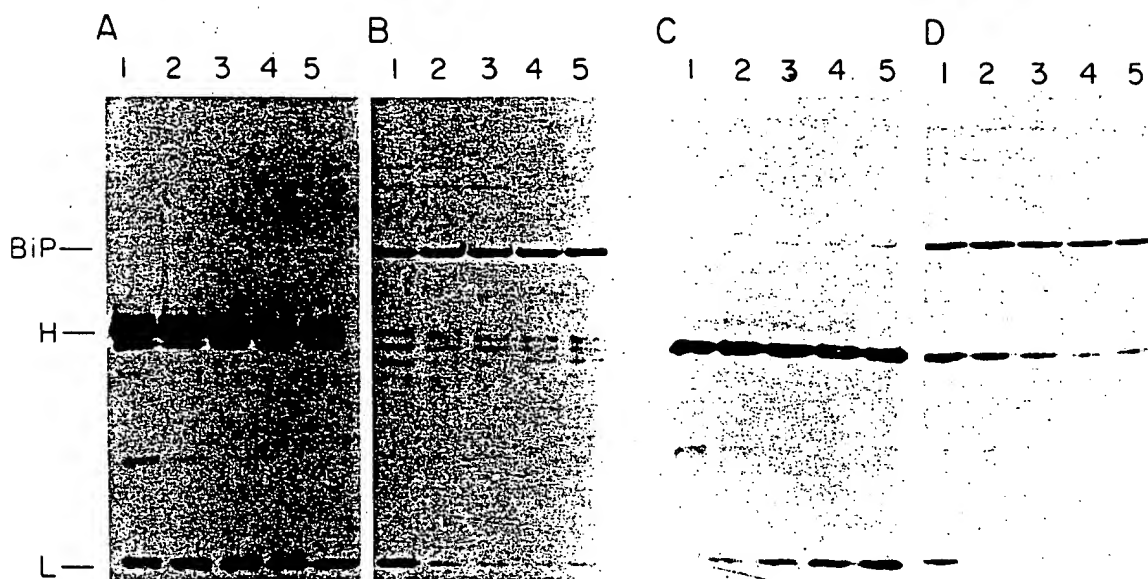


Figure 4. The effect of tunicamycin on immunoglobulin assembly and association with BiP in J558. J558 cells were cultured in the absence (A and B) or presence (C and D) of 2.5 µg/ml tunicamycin 30 min before labeling. Cells were labeled for 10 min with [³⁵S]methionine and then chased in excess cold methionine for 0, 10, 20, 30, and 60 min (lanes 1 to 5, each panel) before lysis in detergent. Lysates were immunoprecipitated with goat anti-α heavy chain antibody, and protein A-Sepharose (A and C) or with monoclonal rat anti-BiP and protein A-Sepharose (B and D). Immunoprecipitated proteins were reduced and resolved on 10% SDS PAGE. Fluoroenhanced gels containing anti-BiP immunoprecipitates (B and D) were exposed to film approximately twice as long as those containing anti-α immunoprecipitates (A and C).

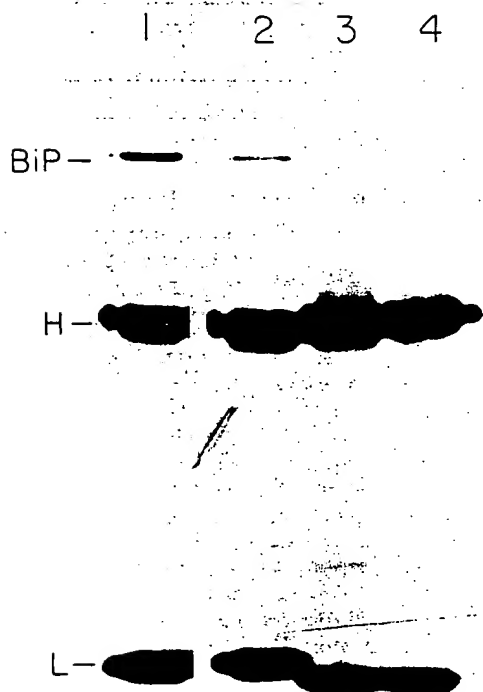


Figure 5. Autoimmune nature of anti-BiP antibody. Rat X Ag8.653 heterohybrid 2-1 cells secreting anti-BiP antibody were labeled for 15 min and then lysed in detergent and immunoprecipitated with protein A-Sepharose (lane 1). Cell-free culture supernatant from 2-1 cells labeled for 15 min and chased 4 h was immunoprecipitated with goat anti-rat IgG coupled to Sepharose 4B (lane 2). Cell-free culture supernatant from rat heterohybridoma 11-11 cells labeled for 15 min and chased for 4 h was immunoprecipitated with goat anti-rat IgG Sepharose (lane 3). Labeled cell free supernatant from 11-11 was

co-precipitated with secreted Ig from the 2-1 heterohybridoma (Fig. 5, lane 2). The co-precipitating protein comigrated on SDS PAGE with intracellular BiP immunoprecipitated from a cell lysate of 2-1 (Fig. 5, lane 1). We believe this to be BiP that has been secreted in an immune complex with a small fraction of the secreted antibody 2-1.

To examine what possible effects such interaction may have on the cellular location of intracellularly immune-complexed BiP, the heterohybridoma, 2-1, and a control heterohybridoma, 11-11, were labeled with [³⁵S]methionine and the cell-free culture supernatants immunoprecipitated with goat anti-rat Ig. Heavy and light chains were secreted from both the 2-1 and 11-11 culture supernatants (Fig. 5, lanes 2 and 3). However, a small amount of an additional protein of *M*_r 78,000 co-precipitated with secreted Ig from the 2-1 heterohybridoma (Fig. 5, lane 2). The co-precipitating protein comigrated on SDS PAGE with intracellular BiP immunoprecipitated from a cell lysate of 2-1 (Fig. 5, lane 1). We believe this to be BiP that has been secreted in an immune complex with a small fraction of the secreted antibody 2-1.

If BiP is membrane-associated or is tightly associated with an intracellular structure, it is probable that antibody immune complexed to BiP would remain within intracellular sites. The co-precipitation of BiP from the culture supernatant of 2-1 suggests that BiP is a soluble protein which when immune complexed to specific antibody can be transported in unison along the secretory pathway and subsequently secreted together with Ig. This demonstrates in unbroken, viable cells

mixed with an equal volume of unlabeled 2-1 supernatant and immunoprecipitated with goat anti-rat Ig Sepharose (lane 4). Immunoprecipitated proteins were reduced and resolved on 10% SDS PAGE.

THIS PAGE BLANK (USPTO)

that BiP and Ig at some time must exist within the same intracellular compartment.

To rule out the possibility that a small amount of BiP leaked from 2-1 cells while in culture and that immune complex formation occurred after secretion of antibody into the culture fluid, we immunoprecipitated unlabeled supernate from 2-1 that was mixed with labeled supernate from 11-11 with goat anti-rat Ig. Precipitated proteins analyzed by SDS PAGE revealed only the precipitation of 11-11 heavy and

light chains; no band migrating with M_r 78,000 was detected (Fig. 5, lane 4). Therefore, the anti-BiP antibody in the unlabeled 2-1 supernatant was unable to immune complex any labeled BiP in the supernatant of 11-11. Immunoprecipitation of cell culture supernatants from other cell lines with anti-BiP antibody has failed to detect extracellular BiP. This suggests that the presence of BiP in the supernatant of 2-1 is due to specific co-transport with anti-BiP antibody and is not a result of nonspecific leakage.

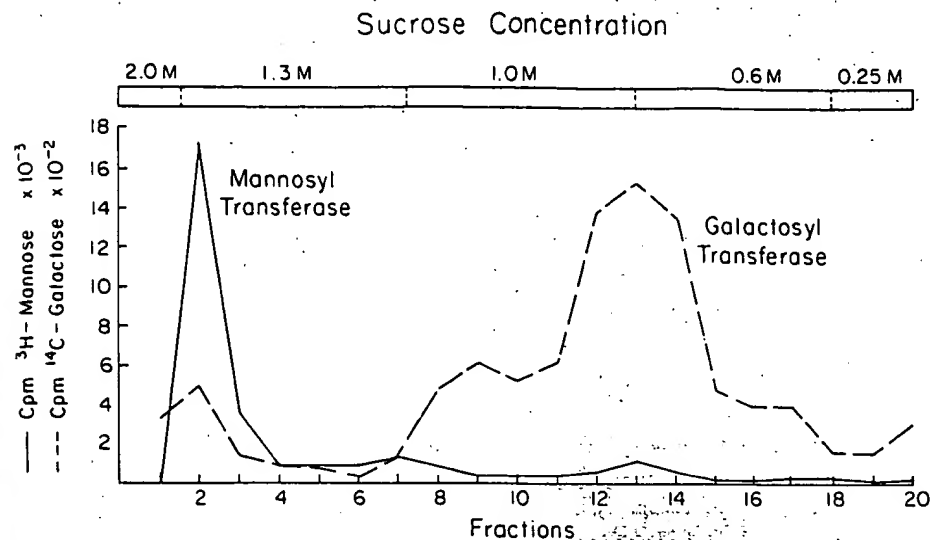
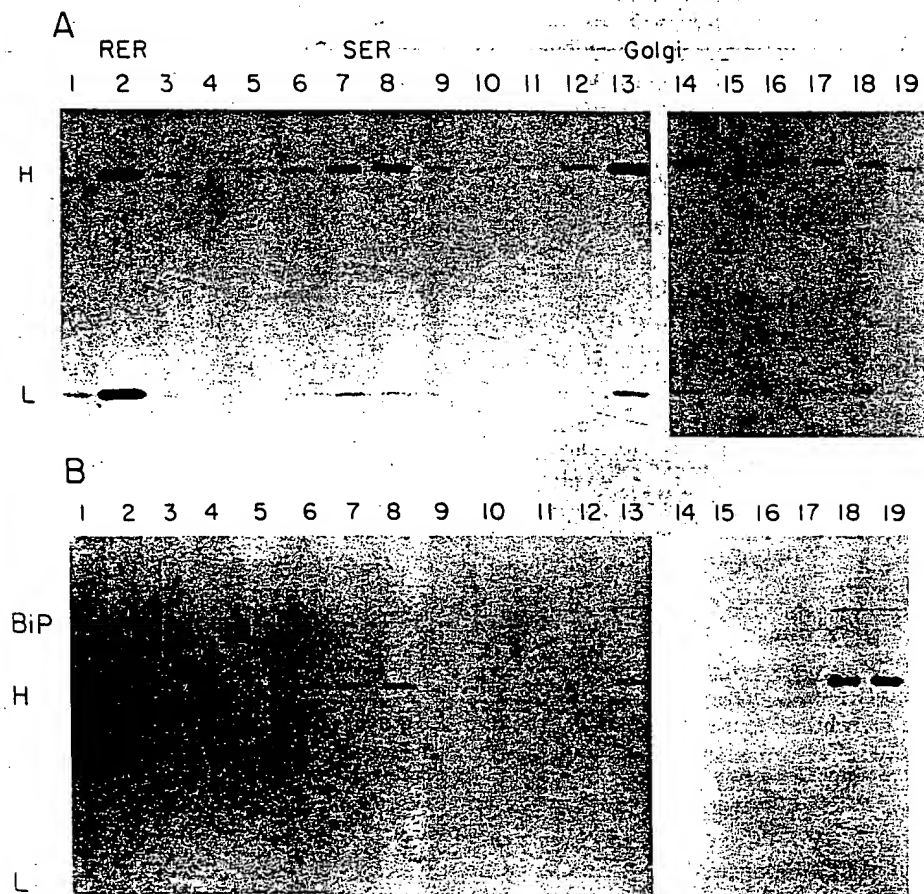


Figure 6. Subcellular fractionation of RD3-2 and Ag8(8). RD3-2 cells were disrupted by nitrogen cavitation and a postnuclear supernatant layered over a discontinuous sucrose gradient. After centrifugation membrane containing fractions were collected from the bottom of the tube. Marker enzymes for RER, mannosyltransferase, and Golgi apparatus: galactosyltransferase were assayed. The marker enzyme profiles were obtained from RD3-2. A similar distribution of marker enzyme activity was obtained from Ag8(8) (not shown). For immunoprecipitation of Ig heavy chains from fractionated membrane vesicles of RD3-2 or Ag8(8), cells were labeled for 15 min and then chased 30 min in the presence of excess cold methionine. Labeled cells were then subjected to nitrogen cavitation and membrane fractions were resolved over a discontinuous sucrose gradient. To each fraction was added an equal volume of 2 \times lysis buffer. Labeled Ig was immunoprecipitated from membrane fractions of RD3-2 (A) or Ag8(8) (B) with goat anti- γ_1 antibody and protein A-Sepharose. Immunoprecipitates were reduced and proteins separated by 10% SDS PAGE.



THIS PAGE BLANK (USPTO)

Intracellular Location of BiP and Immunoglobulin Heavy Chains in Secretory RD3-2 and in Nonsecretory Ag8(8) Hybridomas

The analysis of the association of BiP with Ig heavy chains suggests this interaction occurs in the endoplasmic reticulum. Since heavy chains, in heavy chain-only cell lines, are not secreted and have a prolonged association with BiP it is possible that heavy chains associated with BiP remain within the RER. Resistance of oligosaccharides to endoglycosidase H cleavage has been used as a marker for measuring intracellular transport of many glycoproteins to the Golgi apparatus. However, endoglycosidase H-resistant oligosaccharides cannot be demonstrated among intracellular Ig of many mouse myelomas (15, 41). Thus terminal glycosylation is observed as a co-secretional event within mouse myeloma cells. For this reason we separated membranes of the RER, smooth endoplasmic reticulum (SER), and Golgi apparatus to examine the intracellular location of BiP and associated heavy chains in secreting RD3-2 and nonsecreting Ag8(8).

RD3-2 and Ag8(8) cells were disrupted by nitrogen cavitation and a postnuclear supernatant containing membrane vesicles was fractionated on a discontinuous sucrose gradient. Membranes enriched for RER, SER, and Golgi apparatus were recovered at the three interfaces of the sucrose layers after centrifugation. Fig. 6 shows the location of the marker enzyme mannosyl transferase for RER and galactosyl transferase for the *trans*-Golgi region. Membranes banding at the second interface between fractions 7 and 8 are believed to be enriched for SER. Mannosyltransferase activity can be shifted from fraction 2 to fraction 7 in the presence of EDTA (data not shown). The majority of galactosyltransferase activity is concentrated around fraction 13 (Golgi membrane); however, significant amounts of galactosyltransferase activity are also present in fractions designated RER and SER. The fractions designated RER and SER do not contain mannosyltransferase activity above background, and thus were considered free of RER-derived membranes.

Cells were labeled with [³⁵S]methionine for 15 min and given a 30-min chase before disruption. Membrane fractions from RD3-2 and Ag8(8) were analyzed for presence of Ig and BiP. Membrane vesicles in each fraction were solubilized with lysis buffer and then immunoprecipitated separately with anti- γ_1 heavy chain and anti-BiP antibody. Immunoprecipitated Ig from RD3-2 cells was present in greatest quantities in fractions enriched for RER, SER, and Golgi membranes (Fig. 6A). In addition, Ig was immunoprecipitated from fractions 17, 18, and 19 which represents Ig in the cell disruption buffer that does not sediment with membrane vesicles. This most likely is Ig that has leaked from membrane vesicles during the cell disruption procedure (36). Fractionated membranes from Ag8(8) show a different migrational profile for heavy chains within the sucrose gradient than those of RD3-2 (Fig. 6B). Heavy chains which are associated with BiP are present in greatest amounts in fraction 2 containing RER membranes; only a very small amount of heavy chain is seen in the SER, or Golgi fractions. Densitometer tracings were performed on the autoradiographs (Fig. 6, A and B) to quantitate the relative amounts of heavy chain within each fraction. The results obtained are shown in Table I. In RD3-2, 41% of the γ_1 heavy chains were found in the RER and 38% in the Golgi apparatus, whereas in Ag8(8) 73% of the heavy chains were within

Table I. Subcellular Localization of Intracellular Ig and BiP in RD3-2 and Ag8(8)

Cell type and protein precipitated	Subcellular fraction			
	RER	SER	Golgi apparatus	Cytosol
	%	%	%	%
Ag8(8), γ_1 30-min chase	73	6	3	18
RD3-2, γ_1 30-min chase	41	17	38	4
Ag8(8), BiP 30-min chase	61	2	<2	36
RD3-2, BiP 30-min chase	82	<1	<1	17
RD3-2, BiP 4-h chase	60	3	<2	33

Labeled Ig heavy chains or BiP was quantitated by densitometer tracings. Fractions 1, 2, and 3 were designated RER, fractions 7 and 8 SER, fractions 12, 13, and 14 Golgi apparatus, and fractions 17, 18, and 19 were designated cytosolic or nonmembrane vesicle-associated protein. The percent of total intracellular heavy chain or BiP within each subcellular fraction was determined by dividing the relative densitometer value of a subcellular fraction by the value of the total of all fractions.

the RER and only 3% in the Golgi apparatus. The subcellular fractions from Ag8(8) and RD3-2 were also examined for presence of BiP by immunoprecipitation. Table I shows that BiP is found largely in RER fractions from each cell type, and is not present in other membrane fractions in quantities larger than 3% of total. The presence of BiP in the cytosolic fractions supports the hypothesis that BiP is soluble and not a membrane-associated protein. BiP was also extracted into the aqueous phase during Triton X-114 detergent partitioning (not shown), a procedure which promotes separation of integral membrane proteins into a detergent phase and soluble proteins into the aqueous phase (6). Thus BiP is most likely located within the RER and can be demonstrated to behave as a freely soluble protein.

Discussion

The studies described in this paper focus on the association of the microsomal protein BiP with intracellular Ig heavy chains from secreting (H^+ , L^+) and nonsecreting (H^+ , L^-) mouse B lymphocyte cell lines. A direct correlation exists between light chain addition, BiP displacement and the ability of assembled Ig molecules to be transported to the Golgi apparatus.

The detection of BiP heavy chain complexes in (H^+ , L^+) cell lines would be anticipated to be dependent on both the molar ratio of heavy and light chains and the mode and rate of assembly of the chains within a given cell line. Analysis of a variety of myelomas has shown that heavy and light chains are synthesized in variable ratios, however, the majority have a molar excess of light over heavy chains (2). Thus the detectable amount of free heavy chain associated with BiP may be very small when heavy chain concentration limits assembly. Our ability to readily detect BiP heavy chain complexes in RD3-2 and not in MM60 may reflect a difference in H:L chain ratios as well as the mode of light chain addition. RD3-2 assembles heavy chain dimers (H_2) before light chain addition, whereas in IgM-producing cell lines such as MM60, light chains are added directly to individual heavy chains (5, 30). Consequently, if BiP displacement is mediated solely by light chain addition, heavy chains available for BiP association would exist within RD3-2 until the last light chain is added forming the H_2L_2 tetramer, while in MM60 BiP would be



al

THIS PAGE BLANK (USPTO)

displaced immediately after the light chain is added forming the HL dimer.

In RD3-2 and J558, co-precipitation of BiP with assembling Ig precursors was readily observed when anti-BiP antibody was used as the precipitating agent. However, when assembling heavy chains were precipitated with anti-heavy chain antibody, little or no co-precipitating BiP was observed. It is not known why the BiP heavy chain complex is more readily detectable from the anti-BiP immunoprecipitates. This observation may explain why BiP has not previously been identified to co-precipitate with assembling Ig heavy chains when analyzed with polyclonal anti-heavy chain antibody reagents.

Experiments with J558 revealed that when glycosylation of heavy chains was inhibited with tunicamycin, an increased level of BiP heavy chain complex was observed in pulse-chase experiments. Inhibition of heavy chain glycosylation has been reported to relieve a translational block in heavy chain synthesis. Relief of this block resulted in a higher H/L ratio (4). If tunicamycin relieves a translational block of heavy chain synthesis in J558, the increase in BiP heavy chain complexes observed in the presence of tunicamycin could result from an increase in the intracellular pool of heavy chain. Alternatively, inhibition of glycosylation could prevent normal protein folding resulting in a heavy chain conformation that associates less efficiently with light chain (10, 18).

Subcellular fractionation of Ag8(8) demonstrated that transport of free heavy chains to the Golgi compartment is either blocked or is inefficient when compared to the transport of completely assembled molecules in RD3-2. This finding is in agreement with the differential transport of membrane and secretory μ -chains within the human B-cell line, Daudi. Daudi cells produce both secretory and membrane μ -chains, however only the membrane μ -chains are externalized (10). The secretory heavy chains do not assemble with light chain and remain endoglycosidase H sensitive and thus are blocked from transport to, at least, the medial Golgi region (Hendershot, L., manuscript in preparation). BiP was only immunoprecipitated in significant quantities from the RER membrane-enriched vesicles. This finding demonstrates that BiP is a native component of the RER and maintains its subcellular location despite a continual synthesis and export of newly synthesized secretory protein. Thus, the transport of Ig from RER to Golgi apparatus is a very selective process restricting transport of native RER luminal components. A similar restriction of subcellular location has also been shown for three membrane-associated RER-specific proteins within a mouse myeloma (23).

Intracellular transport of secretory proteins from the RER to the Golgi apparatus has been hypothesized to be mediated by receptors within the endoplasmic reticulum (12, 24, 40). Receptor-mediated intracellular transport would ensure selective transport of only those proteins expressing determinants recognized by such receptors. Our data demonstrating selective transport of assembled Ig directly supports such a hypothesis. In the absence of light chain assembly, signals necessary for transport are not expressed on heavy chains associated with BiP. Transport signals within the heavy chain may be sterically masked by BiP, or BiP may influence the folding of the heavy chain such that proper conformation necessary for transport is not achieved. Light chain addition with disas-

sociation of BiP appears to be necessary for initiating transport. This suggests that BiP may play a regulatory role in the intracellular transport of Ig molecules by preventing the premature transport and possible expression of Ig heavy chains before their assembly with light chain.

Acquisition of proper conformation has been demonstrated to be necessary for the transport of other secretory proteins. In this regard, the transport of retinol-binding protein requires exogenous ligand for exit from the RER (34). One and two amino acid substitutions within secretory proteins can prevent their secretion presumably by altering conformation (44, 45). In addition, the prevention of glycosylation or its processing, which may influence protein folding, have been shown to affect efficiency of intracellular transport of Ig as well as other secretory proteins (22, 25). The signals necessary for directing transport and secretion of Ig molecules as well as other secretory proteins are enigmatic. Secretion of assembled Ig can occur in cells producing mutant heavy chains with deletions of any one of the constant region domains (20). Thus no one domain is necessary for secretion. In contrast to heavy chains, some but not all light chains can be secreted unassembled when light chain synthesis is in excess of heavy chain synthesis or when cells have lost the ability to produce heavy chains (19). These observations may suggest that neither the heavy chain nor light chain in itself is sufficient for signaling transport, but that it is the product of the two chains within the assembled molecule that signals transport. Heavy chains, however, can possess signals sufficient for secretion. In contrast to μ -chains in pre-B cells, heavy chains can be secreted in the absence of light chain synthesis in lymphoproliferative heavy chain disease (35) and in some mutant myelomas (9, 29). Examination of many heavy chain disease proteins has shown that they contain deletions that include portions of the first constant region (CH_1) and/or variable region domains where light chains associate (35). In preliminary experiments, we have been unable to demonstrate association of BiP with heavy chains containing deletions within the CH_1 domain. We speculate that association of nascent heavy chains with BiP requires an intact CH_1 domain.

We have described a novel posttranslational processing event involving the association of unassembled intracellular Ig heavy chains with the microsomal protein BiP. This association occurs early in the maturational process of immunoglobulin assembly within the endoplasmic reticulum. We propose that the association of BiP with unassembled heavy chains may prevent their transport along the secretory pathway by masking expression of transport recognition signals within the heavy chain molecule. We have found BiP to be ubiquitously distributed among cells of different lineages and from the results of preliminary experiments speculate that BiP may function in posttranslational processing and transport of other secretory proteins as well.

We wish to thank Bill Jourdan and Eric Hunter for their comments and suggestions, and Ann Brookshire for her help in preparing the manuscript.

This work has been supported by grants CA 16673, CA 13148, and AI 14782 from the National Institutes of Health. John F. Kearney is the recipient of a Research Career Development Award, AI 00338.

Received for publication 25 November 1985, and in revised form 31 January 1986.

THIS PAGE BLANK (USPTO)

References

- Anderson, D. R., and W. J. Grimes. 1982. Heterogeneity of asparagine-linked oligosaccharides of five glycosylation sites on immunoglobulin M heavy chain from mineral oil plasmacytoma 104E. *J. Biol. Chem.* 257:14858-14864.
- Baumal, R., and M. D. Scharff. 1973. Synthesis, assembly, and secretion of γ -globulin by mouse myeloma cells. *J. Immunol.* 111:448-456.
- Bergman, L. W., and W. M. Kuehl. 1979. Co-translational modification of nascent immunoglobulin heavy and light chains. *J. Supramol. Struct.* 11:9-24.
- Bergman, L. W., E. Harris, and W. M. Kuehl. 1981. Glycosylation causes an apparent block in translation of immunoglobulin heavy chain. *J. Biol. Chem.* 256:701-706.
- Bevan, M. J., R. M. E. Parkhouse, A. R. Williamson, and B. A. Askonas. 1972. Biosynthesis of Immunoglobulin. *Prog. Biophys. Mol. Biol.* 25:131-162.
- Bordier, C. 1981. Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* 256:1604-1607.
- Brew, K., J. H. Shaper, K. W. Olsen, I. P. Trayer, and R. L. Hill. 1975. Cross-linking of the components of lactose synthetase with dimethylpimelimidate. *J. Biol. Chem.* 250:1434-1444.
- Burrows, P. M., M. LeJeune, and J. F. Kearney. 1979. Evidence that murine pre-B cells synthesis μ -heavy chains but no light chains. *Nature (Lond.)* 280:838-841.
- Dackowski, W., and S. L. Morrison. 1981. Two α heavy chain disease proteins with different genomic deletions demonstrate that nonexpressed α heavy chain genes contain methylated bases. *Proc. Natl. Acad. Sci. USA.* 78:7091-7095.
- Dulis, B. H., T. M. Kloppel, H. M. Grey, and R. T. Kubo. 1982. Regulation of catabolism of IgM heavy chains in a B lymphoma cell line. *J. Biol. Chem.* 257:4369-4374.
- Fitting, T., and D. Kabat. 1982. Evidence for a glycoprotein "signal" involved in transport between subcellular organelles. *J. Biol. Chem.* 257:14011-14017.
- Fries, E., L. Gustafsson, and P. A. Peterson. 1984. Four secretory proteins synthesized by hepatocytes are transported from endoplasmic reticulum to Golgi complex at different rates. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:147-152.
- Haas, I. G., and M. Wabl. 1984. Immunoglobulin heavy chain binding protein. *Nature (Lond.)* 306:387-389.
- Hickman, S., and S. Kornfeld. 1978. Effect of tunicamycin on IgM, IgA, and IgG secretion by mouse plasmacytoma cells. *J. Immunol.* 121:990-996.
- Hickman, S., J. L. Theodorakis, J. M. Greco, and P. H. Brown. 1984. Processing of MOPC 315 immunoglobulin A oligosaccharides: evidence for endoplasmic reticulum and trans-Golgi α 1,2-mannosidase activity. *J. Cell Biol.* 98:407-416.
- Kearney, J. F., A. Radbruch, B. Liesegang, and K. Rajewsky. 1979. A new mouse myeloma cell line that has lost immunoglobulin expression but permits the construction of antibody-secreting hybrid cell lines. *J. Immunol.* 123:1548-1550.
- Kearney, J. F., D. E. Briles, and M. J. LeJeune. 1981. A study of immunoglobulin diversity expressed by hybridomas made with pre-B cells and immature B lymphocytes. In *Monoclonal Antibodies and T-cell Hybridomas*. G. J. Hammerling, U. Hammerling, and J. F. Kearney, editors. Elsevier/North-Holland Biomedical Press, 379-387.
- Kleine, R., F. V. Shimakova, V. A. Lapuk, G. V. Vikha, and E. D. Kaverzneva. 1975. Different behavior of immunoglobulin M poor in carbohydrate and native immunoglobulin M during dissociation and reassociation in vitro. *Immunochemistry* 12:825-831.
- Köhler, G., S. C. Howe, and C. Milstein. 1976. Fusion between immunoglobulin-secreting and nonsecreting myeloma cell lines. *Eur. J. Immunol.* 6:292-295.
- Köhler, G., M. J. Potash, H. Lehrach, and M. J. Shulman. 1982. Deletions in immunoglobulin μ chains. *EMBO (Eur. Mol. Biol. Organ.) J.* 1:555-563.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.)* 227:680-685.
- Leavitt, R., S. Schlesinger, and S. Kornfeld. 1977. Impaired intracellular migration and altered solubility of nonglycosylated glycoproteins of Vesicular Stomatitis virus and Sindbis virus. *J. Biol. Chem.* 252:9018-9023.
- Lewis, M. J., S. J. Turco, and M. Green. 1985. Structure and assembly of the endoplasmic reticulum: biosynthetic sorting of endoplasmic reticulum proteins. *J. Biol. Chem.* 260:6926-6931.
- Lodish, H. F., N. Kong, M. Snider, and G. J. A. M. Strous. 1983. Hepatoma secretory proteins migrate from rough endoplasmic reticulum to Golgi at characteristic rates. *Nature (Lond.)* 304:80-83.
- Lodish, H. F., and N. Kong. 1984. Glucose removal from N-linked oligosaccharides is required for efficient maturation of certain secretory glycoproteins from the rough endoplasmic reticulum to the Golgi complex. *J. Cell Biol.* 98:1720-1729.
- Mains, P. E., and C. H. Sibley. 1982. The requirement of light chain for the surface deposition of the heavy chain of immunoglobulin M. *J. Biol. Chem.* 258:5027-5033.
- Maki, R., J. Kearney, C. Paige, and S. Tonegawa. 1980. Immunoglobulin gene rearrangement in immature B cells. *Science (Wash. DC)* 209:1366-1369.
- Morrison, S. L., and M. D. Scharff. 1975. Heavy chain-producing variants of a mouse myeloma cell line. *J. Immunol.* 114:655-659.
- Morrison, S. L. 1978. Murine heavy chain disease. *Eur. J. Immunol.* 8:194-199.
- Parkhouse, R. M. E. 1971. Immunoglobulin M biosynthesis: production of intermediates and excess of light-chain in mouse myeloma MOPC 104E. *Biochem. J.* 123:635-641.
- Parikh, I., S. March, and P. Cuatrecasas. 1974. Topics in the methodology of substitution reactions with agarose. *Methods Enzymol.* 34:77-102.
- Perry, R. P., D. E. Kelly, C. Coleclough, and J. F. Kearney. 1981. Organization and expression of immunoglobulin genes in fetal liver hybridomas. *Proc. Natl. Acad. Sci. USA.* 78:247-251.
- Pollok, B. A., A. S. Bhowan, and J. F. Kearney. 1982. Structural and biological properties of a monoclonal auto-anti-(anti-idiotypic) antibody. *Nature (Lond.)* 299:447-449.
- Ronne, H., C. Ocklind, K. Wiman, L. Rask, B. Obrink, and P. A. Peterson. 1983. Ligand-dependent regulation of intracellular protein transport: effect of vitamin A on the secretion of the retinol-binding protein. *J. Cell Biol.* 96:907-910.
- Seligmann, M., E. Mihaesco, J. L. Preud'homme, F. Danon, and J. C. Brouet. 1979. Heavy chain disease: current findings and concepts. *Immunol. Rev.* 48:145-167.
- Sherr, C. J., I. Schenkein, and J. W. Uhr. 1971. Synthesis and intracellular transport of immunoglobulin in secretory and nonsecretory cells. *Ann. NY Acad. Sci.* 190:250-267.
- Siden, E. J., D. Baltimore, D. Clark, and N. Rosenberg. 1979. Immunoglobulin synthesis by lymphoid cells transformed in vitro by Abelson murine leukemia virus. *Cell* 16:389-396.
- Stohrer, R., and J. F. Kearney. 1983. Fine idiotype analysis of B cell precursors in the T-dependent and T-independent responses to α 1-3 dextran in BALB/c mice. *J. Exp. Med.* 158:2081-2094.
- Stoll, J., A. R. Robbins, and S. S. Krag. 1982. Mutant of Chinese hamster ovary cells with altered mannose 6-phosphate receptor activity is unable to synthesize mannosylphosphoryldolichol. *Proc. Natl. Acad. Sci. USA.* 79:2296-2300.
- Strous, G. J. A. M., and H. F. Lodish. 1980. Intracellular transport of secretory and membrane proteins in hepatoma cells infected by Vesicular Stomatitis virus. *Cell* 22:709-717.
- Tartakoff, A., and P. Vassalli. 1979. Plasma cell immunoglobulin M molecules: their biosynthesis, assembly, and intracellular transport. *J. Cell Biol.* 83:284-299.
- Thorens, B., M.-F. Schulz, and P. Vassalli. 1985. Bone marrow pre-B lymphocytes synthesize immunoglobulin μ chains of membrane type with different properties and intracellular pathways. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:361-368.
- Wabl, M., and C. Steinberg. 1982. A theory of allelic and isotypic exclusion for immunoglobulin genes. *Proc. Natl. Acad. Sci. USA.* 79:6976-6978.
- Wu, G. E., N. Hozumi, and H. Murialdo. 1983. Secretion of a λ 2 immunoglobulin chain is prevented by a single amino acid substitution in its variable region. *Cell* 33:77-83.
- Yoshida, A., J. Lieberman, G. Laima, and C. Ewing. 1976. Molecular abnormality of human α 1-antitrypsin variant (Pi-ZZ) associated with plasma activity deficiency. *Proc. Natl. Acad. Sci. USA.* 73:1324-1328.

THIS PAGE BLANK (USPTO)

Volume 102

Number 5

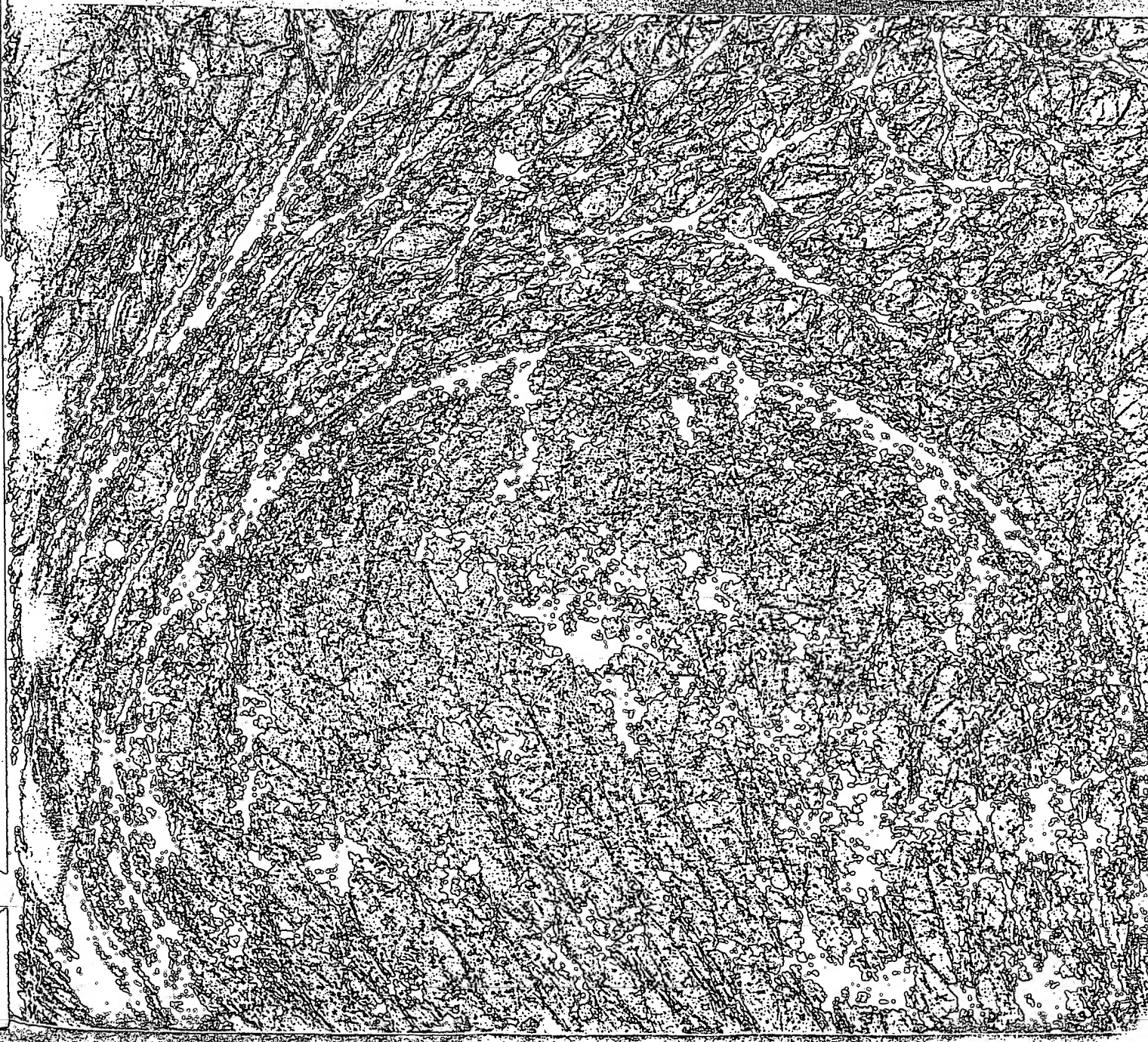
May 1986

Published by
The Rockefeller
University Press
and edited

in cooperation with
The American Society
for Cell Biology

ESTABLISHED 1962
ISSN 0021-9545
MAY 1986

The Journal of Cell Biology



THIS PAGE BLANK (USPTO)

expression on the cell surface. The number of the sizes of the visible human repetitive DNA bands gives us an estimate of the amplified unit of at least 200 kilobase pairs.

Several pieces of evidence lead us to conclude that the structural gene for Leu-2 is amplified in the selected cells expressing high levels of Leu-2: (1) the transfection frequency for Leu-2 using J10-6 DNA is 15–20 times greater than with the same amount of donor JM lymphoma DNA or J10 DNA, (2) double minute chromosomes, known to occur with gene amplification in mouse cells, are present in J10-6 cells, (3) the introduced human or TK DNA integrated in the L-cell genome is amplified 10–50 times based on Southern blots with human repetitive DNA or TK as probes. Spontaneous amplification for the Leu-2 gene in a mouse L-cell transferent may be a fairly common event since at least 4 of 20 independent, FACS isolated, Leu-2 transformed clones showed an increased and unstable pattern of antigen expression. For each of the four clones we were able to select cells with greater antigen expression by FACS selection of the brightest stained cells. Preliminary work indicates that growing cells in HAT medium versus non-HAT medium is more effective for this process.

FACS has been used for selecting cells with other kinds of amplified sequences. Johnston *et al.*⁷ selected for amplification of dihydrofolate reductase (DHFR) in CHO cells by multiple cycles of sorting the brightest cells after staining with fluorescent methotrexate, an inhibitor of DHFR. This group obtained cells that had spontaneously amplified the DHFR gene 50-fold. With suitable fluorescent probes, selection for amplification of other genes may be made in a similar manner.

Cells with amplified copies of a gene coding for a cell-surface antigen could be very useful for cloning this gene. Cloning by cDNA methods could be simplified because the levels of mRNA are almost certainly increased in amplified cells. For example, Caskey *et al.*⁸ isolated a cDNA clone corresponding to the HPRT gene using a neuroblastoma line with 40-fold amplification. Double minutes could be isolated and used for the production of genomic libraries enriched for amplified sequences⁹. Cells with increased amounts of antigen expression should also make protein purification for structural and functional studies of cell-surface antigens easier.

We thank Drs Randal Johnston and David Parks for helpful discussions and Ms Susan Gunther and Ms Maria Chiara Rattazzi for technical assistance. This work was supported by NIH grants GM-17367 and CA-04681. P.K. was supported by NIH fellowship grant AI06343.

Received 24 June; accepted 15 September 1983.

1. Ledbetter, J. A. *et al.* *J. exp. Med.* **153**, 310–323 (1981).
2. Snow, P., Spits, H., deVries, J. & Terhorst, C. *Hybridoma* (in the press).
3. Kavathas, P. & Herzenberg, L. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 524–528 (1983).
4. Kaufman, R. J., Brown, P. C. & Schimke, R. T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5669–5673 (1979).
5. Brown, P. C., Beverley, S. M. & Schimke, R. T. *Molec. Cell Biol.* **1**, 1077–1083 (1983).
6. Peruchio, M., Hanahan, D. & Wigler, M. *Cell* **22**, 309–317 (1980).
7. Johnston, R. N., Beverley, S. M. & Schimke, R. T. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
8. Brennand, J., Chinault, A. C., Konecki, D. S., Melton, D. W. & Caskey, C. T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1950–1954 (1982).
9. George, D. L. & Powers, V. E. *Cell* **24**, 117–123 (1981).
10. Shih, C. & Weinberg, R. A. *Cell* **29**, 161–169 (1982).

Immunoglobulin heavy chain binding protein

Ingrid G. Haas & Matthias Wabl

Friedrich-Miescher-Laboratorium, der Max-Planck-Gesellschaft, Spemannstrasse 37–39, D-7400 Tübingen, FRG

Pre-B lymphocytes, and hybridomas derived from them, synthesize immunoglobulin heavy (IgH) chain in the absence of light (L) chain¹. In the Abelson virus transformed line 18-81, which is representative of the pre-B cell stage, we observed that at least some of the H-chains are bound to a protein other than L-chain. Here we show that the protein (which we term immunoglobulin heavy-chain binding protein, BiP) binds non-covalently to free IgH, but not to IgH associated with IgL.

Some of the 18-81 cells produce μ -chain and others produce γ 2b-chain. After fusion with a myeloma that does not synthesize any Ig chain, hybridomas can be recovered which produce μ - or γ 2b-chain, and some of which also produce κ -chain². Immunoprecipitation of intracellular Ig from the L-chain negative pre-B cell hybridomas resulted in the co-precipitation of a protein migrating with the same electrophoretic mobility as μ -chain in SDS-polyacrylamide gel electrophoresis (PAGE) in reducing conditions. We have previously shown that this protein of molecular weight 78,000 (78K) is not μ -chain². Figure 1a shows this protein from cells of the hybridoma line H6 which produces γ 2b chain but no L-chain (lane 2). In the μ -chain producing hybridoma H13, there is a similar protein, which can be distinguished from μ -chain, because this particular μ -chain migrates slightly faster than normal (Fig. 1a, lane 1). Very small amounts of the 78K protein also co-precipitated with the intracellular (Fig. 1a, lanes 3, 5) but not with the secreted immunoglobulin (lanes 4, 6) from the plasma cell hybridomas Sp2 and PCI 56.0, which produce IgG2b and IgG3, respectively. The 78K protein is not glycosylated (data not shown).

In order to establish whether the 78K protein that co-precipitates with intracellular immunoglobulin is the same in the different cell lines, peptide maps were analysed. The bands of interest were localized by autoradiography, cut out from the dried gels and subjected to limited proteolysis with papain³. In SDS-

PAGE, at least 10 peptides are shared among the 78K proteins from the γ 2b-producing hybridoma H6 (Fig. 1b, lanes 2, 6), from the μ -producing hybridoma H13 (lane 3) and from the plasma cell hybridomas Sp2 and PCI 56.0 (lanes 5 and 7 respectively). The peptide pattern of the γ 2b-chain (lane 1) differs from that of the μ -chain (lane 4) as well as from the peptide pattern of the 78K protein. From these results we conclude that the 78K protein is similar in the different cell lines.

The co-precipitation of the 78K protein with IgH chain could have been due to a cross-reactivity of the antisera used in these experiments. We have excluded this possibility by carrying out the precipitation with *Staphylococcus aureus* in the absence of any antiserum. *S. aureus* binds the Fc portion of mouse γ 2b-chain but not of μ -chain. The 78K protein does not bind to *S. aureus*; that is, no protein is precipitated in a cell lysate of μ -producing cells with *S. aureus* alone (Fig. 2, lane 3). However, on addition of anti- μ antiserum both the 78K protein and μ -chain are precipitated (Fig. 2, lane 5). Since the addition of *S. aureus* alone is sufficient to precipitate both the IgH chain and the 78K protein in γ 2b-synthesizing cells (Fig. 2, lane 6), the co-precipitation must be due to a binding of this protein to the γ 2b-chain. An explanation involving a binding of this protein to the *S. aureus* immunoglobulin complex can be excluded because (1) anti- κ antiserum together with *S. aureus* did not precipitate the 78K protein in the μ -producing hybridoma H13 (Fig. 2, lane 4); (2) the 78K protein was precipitated by anti-H chain antiserum together with a second antibody; (3) the precipitation of H chain and BiP was inhibited by the addition of unlabelled immunoglobulin of the same H-chain class, but not by the addition of immunoglobulin of other classes. Since the 78K protein thus binds to the γ 2b-chain, but also to the H-chain of other classes, we propose that this protein be called immunoglobulin heavy chain binding protein (BiP).

It is possible that the binding of BiP to H-chain is an artefact of the cell lysis and protein precipitation procedure—if BiP and H-chain were present in different cellular compartments the binding would occur only when the cells are lysed with detergent. To test this we analysed subcellular fractions obtained in the absence of detergent. BiP co-precipitated with H-chain in the microsomal fraction and in the soluble cytoplasm (Fig. 3a, lanes 2, 3, Fig. 3b, lanes 2, 3). Although the ratio of H-chain to BiP

THIS PAGE BLANK (USPTO)

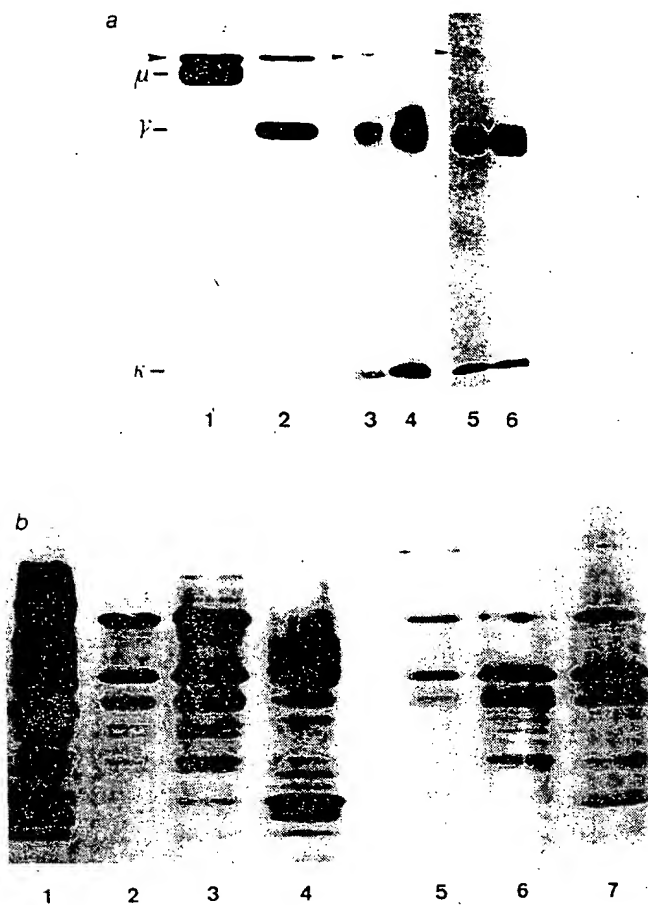


Fig. 1 A 78K protein that is co-precipitated with the intracellular immunoglobulin of various hybridomas. **a**, 2×10^6 cells were cultured in the presence of ^{35}S -methionine, lysed in 0.5% non-ionic detergent and the intracellular immunoglobulin was precipitated as described previously². Secreted immunoglobulin was precipitated from the labelled supernatant of the cells. Analysis of the precipitated material was performed by electrophoresis in 10% SDS-polyacrylamide gels. Lane 1, pre-B-cell hybridoma H13-synthesizing μ -chain precipitated with anti- μ ; lane 2, pre-B-cell hybridoma H6 synthesizing γ 2b-chain precipitated with *S. aureus*; lanes 3, 4, plasma cell hybridoma Sp2 synthesizing IgG2b, intracellular (lane 3) and secreted (lane 4); lanes 5, 6, plasma cell hybridoma PCI 56.0 synthesizing IgG3, intracellular (lane 5) and secreted (lane 6); lanes 3–6, precipitation by *S. aureus* alone. **b**, Limited proteolysis of proteins was performed as described in the text. The ^{35}S -methionine-containing peptides were separated in 20% SDS-polyacrylamide gels and detected by fluorography. γ 2b-chain of H6 (lane 1), 78K protein of H6 (lanes 2, 5), 78K protein of H13 (lane 3), μ -chain of H13 (lane 4), 78K protein of Sp2 (lane 5), 78K protein of PCI 56.0 (lane 7).

varies greatly in different precipitations from whole cell lysates (high in Fig. 3a, b, lane 1, lower in Fig. 1), in the cytoplasm and microsomal fraction the ratio of H:BiP is approximately 1:1, at least when comparing Coomassie blue staining intensity (Fig. 3a, lanes 2, 3). BiP incorporates less ^{35}S -labelled methionine than the γ 2b-chain, which results in a higher H:BiP ratio after autoradiography of the same gel (Fig. 3b). Thus BiP is bound to H-chain when there is no detergent, and each fractionation step caused increased binding of BiP to H-chain. This suggests to us that BiP is associated with H-chain *in situ*, but this point remains to be proven formally.

BiP is bound noncovalently to H-chain in cells of the pre-B-cell line. Its expression in the pre-B-cell hybridoma is not induced by the fusion event. These conclusions are drawn from SDS-PAGE analysis in non-reducing conditions (Fig. 4a). The intracellular μ -chains of a subclone of the 18-81 cell line migrate

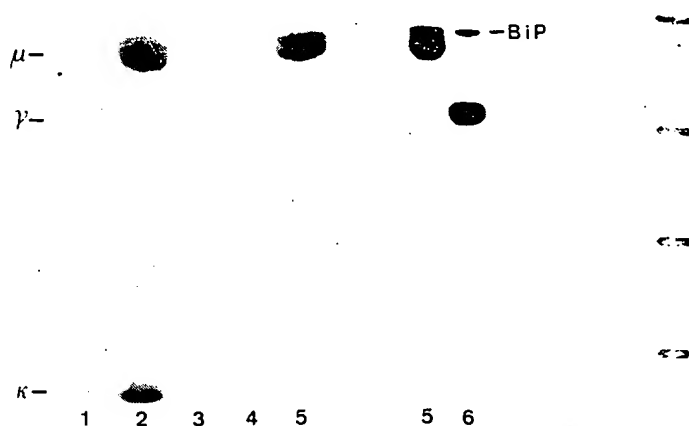


Fig. 2 The 78K protein binds to IgH chain. Experimental procedure was as for Fig. 1a. Precipitation by *S. aureus* was without antiserum in H11, synthesizing μ , κ (lane 1), in H13 (lane 3), and H33-11-2 (lane 6); or by *S. aureus* with antiserum to κ -chain in H11 (lane 2), in H13 (lane 4) or with antiserum to μ -chain in H13 (lane 5).

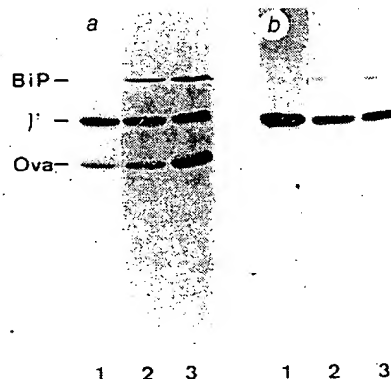


Fig. 3 Recovery of H-chain and BiP in the same subcellular compartments. **a**, Subcellular fractionation of γ 2b-chain synthesizing pre-B-cell hybridoma H33-2 was performed as follows: 2×10^6 ^{35}S -methionine labelled cells were added to 10^8 non-labelled cells, and washed in 0.25 M sucrose. Lysis and immunoglobulin precipitation were performed as described in Fig. 1 with 1/10 of these cells (lane 1). The rest of the cells were homogenized in 1.5 ml sucrose, layered on 2 ml 0.34 M sucrose, centrifuged for 10 min at 600g to remove the nuclei and unbroken cells. After removal of the mitochondrial fraction (10 min at 5,000g), the soluble cytoplasmic proteins were separated from the microsomal fraction by centrifugation at 54,000g for 1 h. The pellet of the microsomal fraction was solubilized with non-ionic detergent (0.5% Triton X-100). Precipitation was performed by addition of *S. aureus* to the microsomal fraction (lane 2) and to the soluble cytoplasmic proteins (lane 3). The cytoplasmic proteins thus never came in contact with detergent. The gel was stained with Coomassie blue; Ova: ovalbumin. **b**, Autoradiography of the dried gel shown in **a**.

mainly as dimers (lane 3; the broadness of this band is presumably due to varying degrees of glycosylation of the μ -chains). Some of the μ -chains migrate as monomers. BiP migrates differently. It is resolved into two bands with a difference in apparent molecular weight of $\sim 1,000$. The same gel pattern is obtained with the μ -producing pre-B-cell hybridoma H13 (lane 2). BiP is also resolved into the same two bands in a γ 2b-producing pre-B-cell hybridoma (lane 4). Both of these bands give rise to a 78K protein when rerun in reducing conditions (Fig. 4b). All other bands contain only γ 2b-chain.

When the cells of the 18-81 cell line are fused with a myeloma which does not synthesize any immunoglobulin chain, some

THIS PAGE BLANK (USPTO)

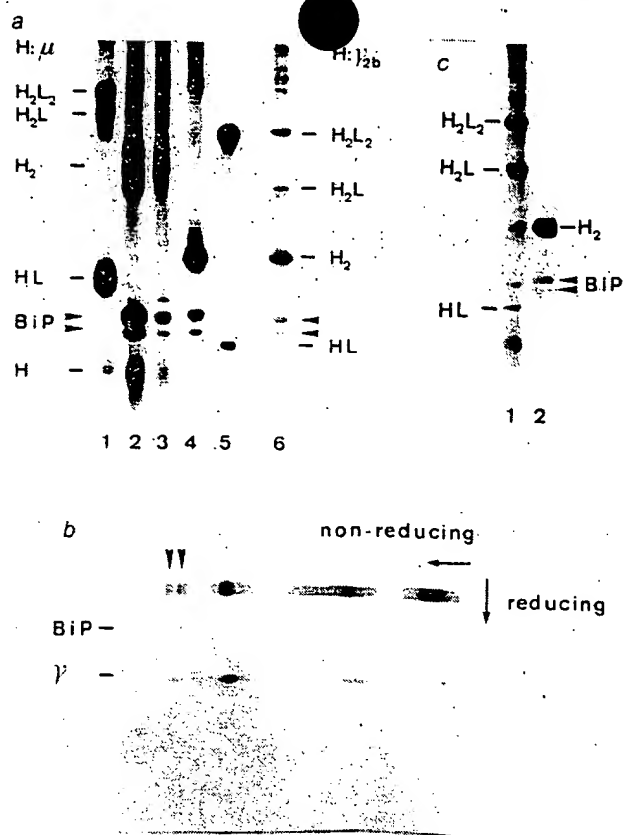


Fig. 4 Noncovalent binding of BiP to IgH-chain in pre-B cells and in hybridomas, but only when H-chain is not associated with L-chain. **a**, Cell labelling and immunoprecipitation as in Fig. 1. Immunoglobulin precipitates were analysed in non-reducing conditions in 5% SDS-polyacrylamide gels: H11, synthesizing μ , κ (lane 1), H13, synthesizing μ -chain (lane 2), 18-81 A33, synthesizing μ -chain (lane 3), H32-2, synthesizing $\gamma 2b$ -chain (lane 4), H32, synthesizing $\gamma 2b$ - κ (lane 5), H33-11-2, synthesizing μ , $\gamma 2b$ and κ (lane 6, precipitation of μ not included). **b**, A $\gamma 2b$ -BiP complex was first analysed in non-reducing conditions as described in **a**. Lane 4 of **a** was cut out from the dried gel, layered on top of a 10% SDS-polyacrylamide gel and rerun in reducing conditions. The $\gamma 2b$ H-chain is found to smear throughout the lane of the nonreducing gel. **c**, Analysis of consecutively precipitated immunoglobulin from H32-1, synthesizing $\gamma 2b$ - κ , in non-reducing conditions. Immunoprecipitation was performed with anti- κ antiserum coupled to Sepharose 4B (lane 1), the remaining immunoglobulin was precipitated by the addition of *S. aureus* (lane 2).

hybridomas can be recovered that produce κ -chain. In these hybridomas there is almost no BiP bound to H-chain (Fig. 4a, lane 1, shows a hybridoma producing $\mu\kappa$, and lane 5 shows a hybridoma synthesizing $\gamma 2b$ - κ), although the H-chain is identical to the μ - or $\gamma 2b$ -chain⁴ associated with BiP in the absence of κ -chain. Two-dimensional gel analysis of total cellular proteins revealed that BiP is still present in these κ -producing cells (data not shown). In hybridoma H33-11-2, not all of the IgH chains are associated with L-chains, and BiP is still synthesized and is co-precipitated in this κ -producing cell (Fig. 4, lane 6). Precipitation with anti- κ antiserum coupled to Sepharose 4B shows the occurrence of mixed molecules (H+H+L+BiP) (Fig. 4c, lane 1). Subsequent precipitation with *S. aureus* alone recovered H-chain dimers associated with BiP (Fig. 4c, lane 2). These findings can be explained in one of the following ways: (1) The binding of BiP might be prevented through steric hindrance by κ -chain; (2) a configuration change of the H-chain might occur when it is covalently bound to L-chain so that the binding site for BiP is altered; (3) BiP and L-chain might compete for the same binding site. Whatever the explanation,

the fact is that the binding of BiP to H-chain is abolished when κ -chain binds to H-chain; this suggests that BiP is involved in the regulation of immunoglobulin chain synthesis—perhaps in the control of allelic and isotypic exclusion for immunoglobulin genes⁵.

Received 4 May; accepted 27 September 1983.

1. Burrows, P. D., LeJeune, M. & Kearney, J. F. *Nature* **280**, 838-841 (1979).
2. Burrows, P. D., Beck, G. B. & Wabl, M. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 564-568 (1981).
3. Cleveland, D. W., Fischer, S. G., Kirchner, M. W. & Lämml, U. K. *J. biol. Chem.* **252**, 1102-1106 (1977).
4. Burrows, P. D., Beck-Engeser, G. B. & Wabl, M. R. *Nature* (in the press).
5. Wabl, M. R. & Steinberg, C. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6976-6978 (1982).

All metacyclic variable antigen types of *Trypanosoma congolense* identified using monoclonal antibodies

J. S. Crowe, J. D. Barry, A. G. Luckins*, C. A. Ross* & K. Vickerman

Department of Zoology, University of Glasgow, Glasgow G12 8QQ, UK

* Centre for Tropical Veterinary Medicine, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK

Vaccination against the tsetse-borne trypanosomiasis has proved impossible because of the trypanosome's ability to generate a seemingly inexhaustible number of variable antigen types in the blood or tissues of the host¹. Each variable antigen is a glycoprotein which forms a surface coat on the trypanosome and each glycoprotein is the product of a single gene^{2,3}. The full repertoire of such antigens has not been identified for any trypanosome serodeme (genotype) as yet, but the number of genes coding for variable antigen glycoproteins is estimated to be between 100 and 1,000^{4,5}. We have previously postulated that for *Trypanosoma brucei* the antigen repertoire of the infective metacyclic stage trypanosomes inoculated by the tsetse fly may be considerably smaller than that expressed in the mammalian host⁶⁻⁸. If this is so then protection against infection by the vector becomes an easier proposition, but the actual scale of the metacyclic repertoire is also unknown. We present here evidence that the metacyclic repertoire of a stock of *T. congolense*, the most important of the pathogenic cattle trypanosomes, is limited to 12 variable antigen types.

T. congolense infected tsetse flies produce very few metacyclic trypanosomes and this inhibits examination of metacyclic variable antigen type (M-VAT) heterogeneity using fly-derived metacyclics. Recently, however, a culture system for generating *T. congolense* metacyclics *in vitro* has been developed⁹. Cultures are initiated with trypanosomes isolated from either the proboscis or proventriculus of infected *Glossina morsitans* and 2×10^6 metacyclics can be harvested daily from the supernatant of a mature culture; metacyclics are separated from their uncoated epimastigote predecessors by ion exchange chromatography¹⁰. We have utilized such metacyclics obtained from an uncloned West African stock, TREU 1290, to produce monoclonal antibodies against M-VATs. BALB/c mice were infected using 10^6 live culture metacyclics (UM1 population, Fig. 1). The mice were drug treated on day 5 with 10 mg per kg Berenil and injected again intravenously with 10^6 metacyclics on day 14. Spleen cells were removed 3 days later and fused with the myeloma cell line P3-X63-Ag8-653^{11,12}.

Twenty-five monoclonal antibodies (mAbs) were isolated and these labelled various subpopulations (VATs), of TREU 1290 metacyclics (UM1) using the indirect immunofluorescence technique on acetone-fixed trypanosomes. Twenty of these mAbs labelled exposed epitopes on live metacyclics. The percentage of the population labelled by individual mAbs varied from day

THIS PAGE BLANK (USPTO)

nature

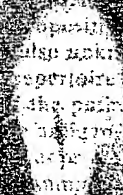
INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

Volume 306: No 5941: 24-30 November 1983: £180. \$4.50

Author Index

CIRCULATION DESK

WIND TUNNEL MODEL FOR FOREST FIRES



306 211 — 300 November 1983

NM029726240EC83 34123006P78 47
COPY 1
STANFORD UNIVERSITY
GOETH LIBRARY - SERIALS
STANFORD
CA 94305
DLR

THIS PAGE BLANK (USPTO)

Immunoglobulin heavy chain toxicity in plasma cells is neutralized by fusion to pre-B cells

(immunoglobulin chain loss/isoelectric focusing of heavy chain)

INGRID C. HAAS* AND MATTHIAS R. WABL*†

*Friedrich-Miescher-Laboratorium der Max-Planck-Gesellschaft, Spemannstrasse 37-39, D-7400 Tübingen, Federal Republic of Germany; and †Basel Institute for Immunology, Grenzacherstrasse 487, CH-4038 Basel, Switzerland

Communicated by Niels K. Jerne, July 9, 1984

ABSTRACT A plasma cell hybridoma frequently loses its immunoglobulin heavy (H) chain spontaneously but rarely its production of its light (L) chain lost. Upon fusion to a pre-B-cell hybridoma that produces no Ig chain, the L chain is frequently lost. In cells without the L chain the H chain, which is derived from the plasma cell, is not chemically modified. Our results indicate that, in pre-B cells, but not in plasma cells, there must be a mechanism that neutralizes the toxic effect of free H chain.

Myelomas and hybridomas derived from plasma cells secrete immunoglobulin consisting of heavy (H) and light (L) chains. From such cell lines, subclones that have lost H chain expression can be recovered, but they still secrete the L chain (1, 2). The L chain can then be lost at the same frequency as the H chain. On the other hand, it seems to be very difficult to recover cells that have lost L chain expression but that still synthesize the H chain (1-3) except when the cells are mutagenized (4) or express a mutant H chain (5-7). These observations have led to the view that the free H chain is toxic to the cells (2, 3). However, cells of an earlier differentiation stage, pre-B cells, synthesize intracellular H chain in the absence of the L chain (8). To explain this difference, several possibilities have been proposed: (i) the rate of H chain synthesis in pre-B cells is too low to damage the cell (9); (ii) the pre-B cell synthesizes a different H chain that is not toxic to the cell (2); or (iii) in pre-B cells, there is a special protein that neutralizes the toxic effect of the free H chain (10, 11).

Here we report that hybridomas derived from pre-B cells are not different from plasma cell hybridomas with respect to their rate of H chain synthesis and steady-state level of H chain. However, they do not synthesize the L chain. In consequence, we wanted to answer the question whether there are pre-B-cell hybridomas that can synthesize a free H chain, which has been shown to be toxic in plasma cell hybridomas.

MATERIALS AND METHODS

Cell Lines. Sp2 and GK14.1 were established and provided by G. Köhler (Basel). The cell lines are derived from fusions between spleen cells and a myeloma, X63 Ag8, and synthesize IgG2b. Sp2.0 is an azaguanine-resistant subclone of Sp2 and has lost Ig expression. Cell lines H32-21, H32-3, H32-8, and H6 are derived from fusions between γ 2b-synthesizing subclones of the Abelson virus-transformed pre-B-cell line 18-81 and X63 Ag8653 (12). Clone H62 is a subclone of H6 and synthesizes no Ig chain. NORA hybridomas were made by fusion of Sp2 HL Ag14 and H62. SPSP hybridomas are derived from a fusion between Sp2.0 and Sp2. Cell fusion

was carried out as described (12). The genealogy of the various hybridomas is given in Fig. 1.

Isolation of Subclones with Ig Chain Loss. Soft agar cloning and antiserum overlay was carried out according to the method of Coffino and Scharff (1). When an antiserum against IgG2b (γ 2b.k) was used, about 1% of the cells formed colonies without precipitation. These clones were isolated, and their Ig expression was analyzed by immunofluorescence. Clones of interest were grown in mass culture and further analyzed by immunoprecipitation and electrophoresis.

NaDodSO₄/Polyacrylamide Electrophoresis. Cell labeling, immunoprecipitation, and NaDodSO₄/polyacrylamide gel electrophoresis were carried out as described (12). Pulse labeling (30 min) was carried out by the addition of [³⁵S]methionine to cells that had been incubated in methionine-free select medium/10% dialyzed fetal calf serum for 1 hr. The amount of [³⁵S]methionine incorporated into the precipitated proteins was measured after gel fractionation by scintillation assay of the solubilized gel slices.

Isoelectric Focusing. [³⁵S]Methionine-labeled Ig precipitates were dissolved in 9.5 M urea/2% Nonidet P-40/2% Ampholine (pH 5-11)/5% 2-mercaptoethanol and applied to isoelectrofocusing slab gels. The gel composition was according to O'Farrell (13). Electrophoresis was carried out for 1 hr at 250 V, for 12 hr at 400 V, and finally for 1 hr at 800 V. The proteins were visualized by fluorography.

Immunofluorescence. The purification and fluorochrome conjugation of goat antibodies specific for mouse H chain isotypes and the methods for immunofluorescence detection of intracellular Ig have been described (14).

RESULTS

The Rate of Ig H Chain Synthesis and the Steady-State Level of the H Chain in Plasma Cell- and In Pre-B Cell-Derived Hybridomas Are of the Same Order of Magnitude. The pre-B cell is characterized by the synthesis of intracellular H chain in the absence of the L chain. The free H chain is toxic in plasma cells, but pre-B cells may survive because they synthesize the H chain in small amounts (15, 16) that would not damage the cell. To test this hypothesis, we increased the amount of free H chain present in pre-B cells by fusion to a myeloma. The steady-state level of RNA specific for the H chain is the same in pre-B-cell hybridomas and in plasma cell hybridomas (16).

By measuring the short-term incorporation of radiolabeled methionine into the H chain (Fig. 2), we have compared the rate of Ig H chain synthesis in pre-B-cell-derived hybridomas with that of plasma cell-derived hybridomas. The amount of radioactivity incorporated into the various H chains was of the same order of magnitude. Also, the steady-state level of the H chain, as determined by long-term label-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: H chain and L chain, heavy and light chain, respectively; of Ig; BiP, H chain binding protein.

THIS PAGE BLANK (USPTO)

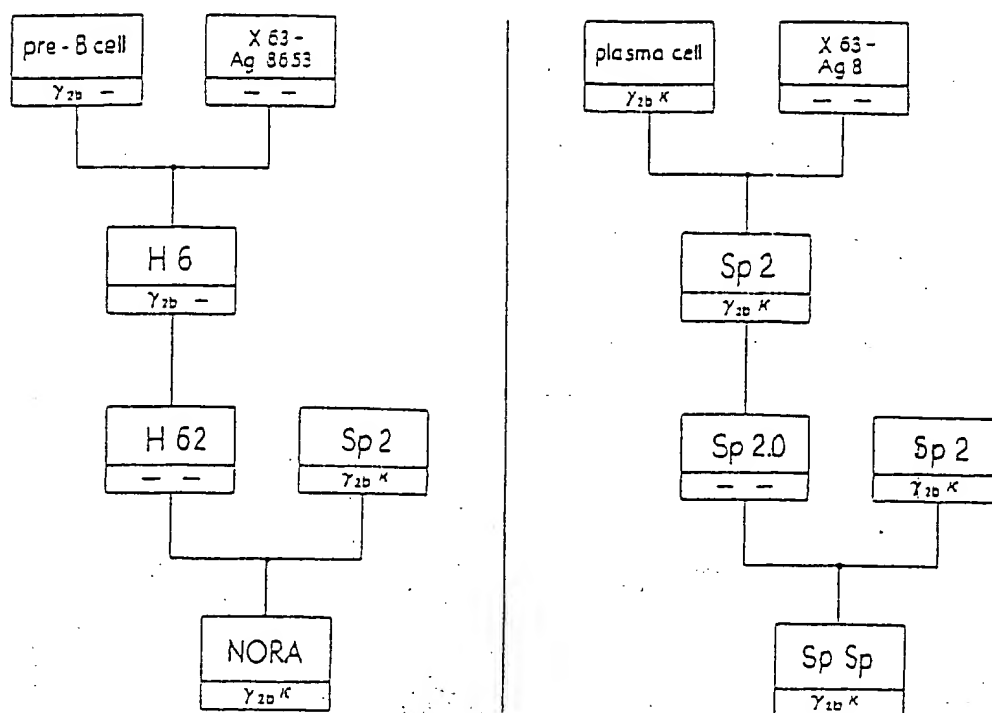


FIG. 1. Genealogy of selected hybridomas derived from a pre-B cell or from a plasma cell, or from both. The kind of Ig chain synthesized is given below the clone designation.

ing (Fig. 3) and by immunofluorescence intensity (data not shown), was not different. This does not formally exclude that pre-B cells survive the expression of free H chain because of a low rate of synthesis, but there would still remain the question of why the H chain is not toxic for pre-B-cell-derived hybridomas as it is for plasma cell-derived hybridomas.

A Given H Chain Is Toxic in Plasma Cells but Not in Pre-B Cells. Because pre-B-cell-derived hybridomas can survive high levels of the H chain, we wanted to know whether they are able to survive the expression of a H chain, the toxicity of which has been demonstrated in plasma cells. For that purpose, we fused a plasma cell hybridoma to a pre-B-cell

hybridoma that had lost its own H chain expression.

The Sp2 cell line is derived from fusion of a plasma cell with a myeloma (Fig. 1). It synthesizes both H and L chain and exhibits the H chain toxicity phenomenon (2). We confirmed this by recovering Ig chain-loss variants according to the method of Coffino and Scharff (1). Of 74 subclones, 69 expressed no H chain, one expressed no L chain (that is, expressed H chain only), and 4 did not express any Ig chain at all (Table 1). We then fused the Sp2 cell line to a pre-B-cell-derived hybridoma, H62, that had lost the expression of its own Ig chain (Fig. 1). The resulting hybrid cell line NORA 4 secreted H and L chain from the Sp2 parent line (Figs. 3 and 4). The NORA cell line tolerates the H chain in the ab-

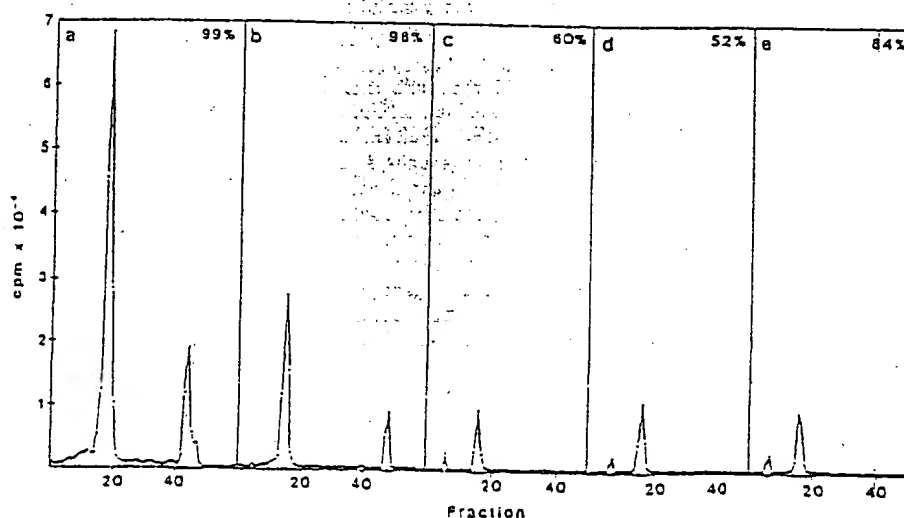


FIG. 2. Amount of radioactivity incorporated into immunoprecipitated proteins after separation on polyacrylamide gel. GK14.1 (a) and Sp2 (b) are plasma cell-derived hybridomas synthesizing both H and L chain; H32-21 (c), H32-3 (d), and H32-8 (e) are pre-B-cell-derived hybridomas synthesizing a H chain that is associated with BiP (11). The indicated percentages of cells synthesizing Ig were determined by immunofluorescence.

THIS PAGE BLANK (USPTO)

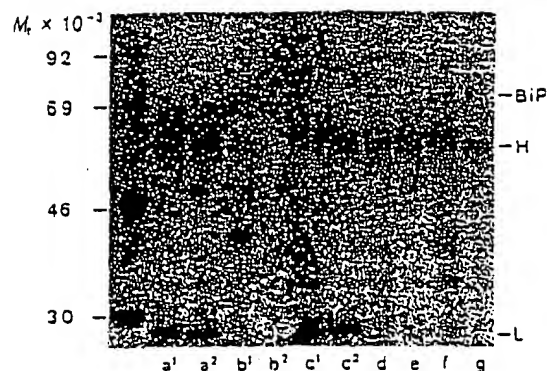


FIG. 3. NaDodSO₄/polyacrylamide gel electrophoresis analysis of the Igs of various hybridomas synthesizing both H and L chain or synthesizing the H chain only. Sp2 (lanes a) and NORA 4 (lanes c) secrete both H and L chain; SPSP 1.55 (lanes b), NORA 4.2 (lane d), NORA 4.16 (lane e), NORA 4.8.20 (lane f), and the pre-B-cell hybridoma H 61 (lane g) synthesize the H chain without the L chain but do not secrete it.

sence of the L chain. Ten out of 55 subclones of NORA 4 with Ig chain loss synthesized Ig H chain without the L chain, 42 synthesized the L chain only, and 3 did not synthesize any Ig chain at all (Table 1). A subclone of NORA 4, NORA 4.8, that synthesizes both H and L chain, showed the same chain loss distribution (Table 1). In no instance was the H chain secreted. As a control to account for the increased chromosome numbers of the NORA 4 hybrid as compared with Sp2, we fused the Sp2 cell line to a nonproducing subclone, Sp2.0, generating SPSP hybridomas (Fig. 1). From these hybridomas, subclones synthesizing free H chain should be rare. Indeed, of 125 subclones with Ig chain loss, we recovered only 1 subclone expressing free H chain (Table 1). Of these subclones, 122 had lost the H chain and two had lost both Ig chains. Comparison of the H chain of hybridomas synthesizing both H and L chain or the H chain alone revealed no difference in size (Fig. 3) nor in isoelectric focusing behavior (Fig. 4). Since in this case, the H chain is one that is known to be toxic, we conclude that in pre-B cells there is a mechanism neutralizing the toxic effect of free H chains.

DISCUSSION

Why are free Ig H chains toxic in plasma cells but not in pre-B cells? From the plasmacytoma MOPC 21, a variant could be recovered that synthesizes only the H chain. The mutant H chain of this variant lacks the last 67 carboxyl-terminal amino acids and forms polymers of at least 20 H chains (9). This polymerization is probably due to free sulfhydryl (SH)

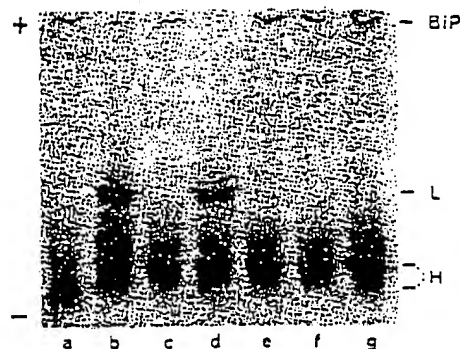


FIG. 4. Isoelectric focusing analysis of intracellular Ig of various hybridomas synthesizing both H and L chain or synthesizing H chain only. Lanes: a, pre-B-cell hybridoma H 61 (γ2b); b, Sp2 (γ2b,κ); c, SPSP 1.55 (γ2b); d, NORA 4 (γ2b,κ); e, NORA 4.2 (γ2b); f, NORA 4.16 (γ2b); g, NORA 4.8.20 (γ2b). The γ2b chain of H 61 has a variable region different from the one of the Sp2 cells and their hybridomas.

groups on the H chain that would normally form the L-H bridge. The normal H chain probably could also polymerize in the absence of L chains but would form much larger insoluble complexes that would damage the cell. Polymerization of L chains cannot occur because no additional free SH groups are available once the L chain has formed a dimer. It is of interest that H chains are predominantly found as monomers or dimers in pre-B cells (11), as well as in the H chain-synthesizing subclones of the NORA 4 hybrid line (data not shown). Thus, neutralization of the toxic effect of the free H chain may be achieved by the prevention of H chain polymerization—for example, by an enzyme that alters the reactive SH groups of the Ig H chain. Since we did not find any difference in the isoelectric focusing pattern of the intracellular H chains of Sp2, NORA 4, and of those subclones synthesizing no L chain (Fig. 4), one can postulate a protein that either rapidly degrades accumulating free H chain or protects the reactive SH groups without chemical modification. We have already described a protein (heavy chain-binding protein, BiP) that binds to Ig H chains not associated with the L chain (11). In all NORA subclones that have lost the L chain (some of them are shown in Fig. 3), the H chain is associated with the BiP. If the BiP is neutralizing H chain toxicity it should be less active in plasma cells than in pre-B cells.

Whatever the mechanism for neutralization of H chain toxicity, our results demand an explanation of why H chain synthesis precedes L chain synthesis in B-cell ontogeny (10).

We thank Drs. J. Johnson (Munich) and C. Steinberg (Basel) for discussions.

Table 1. Ig chain loss pattern in various hybridomas

Hybridoma	No. of subclones		
	H chain lost	L chain lost	H + L chain lost
Sp2	57	0	4
Sp2.68	12	1	0
Total	69	1	4
NORA 4	42	10	3
NORA 4.8	42	12	0
Total	84	22	3
SPSP 1	59	1	2
SPSP 2	63	0	0
Total	122	1	2

- Coffino, P. & Scharff, M. D. (1971) *Proc. Natl. Acad. Sci. USA* 68, 219-223.
- Köhler, G. (1980) *Proc. Natl. Acad. Sci. USA* 77, 2197-2199.
- Wilde, C. D. & Milstein, C. (1980) *Eur. J. Immunol.* 10, 462-467.
- Morrison, S. L. & Scharff, M. D. (1975) *J. Immunol.* 114, 655-659.
- Morrison, S. L. (1978) *Eur. J. Immunol.* 8, 194-199.
- Sunenshein, C. E., Sackevitz, M., Siebert, G. R. & Gefter, M. L. (1978) *J. Exp. Med.* 148, 301-312.
- Cowan, N. J., Secher, D. S. & Milstein, C. (1974) *J. Mol. Biol.* 90, 691-701.
- Burrows, P. D., LeJeune, M. & Kearney, J. F. (1979) *Nature (London)* 280, 838-841.
- Argon, Y., Burrone, O. R. & Milstein, C. (1983) *Eur. J. Immunol.* 13, 301-305.

THIS PAGE BLANK (USPTO)

10. Wabl, M. & Steinberg, C. (1982) *Proc. Natl. Acad. Sci. USA* 79, 6976-6978.
11. Haas, I. G. & Wabl, M. (1983) *Nature (London)* 306, 387-389.
12. Burrows, P. D., Beck, G. B. & Wabl, M. R. (1981) *Proc. Natl. Acad. Sci. USA* 78, 564-568.
13. O'Farrell, P. H. (1975) *J. Biol. Chem.* 250, 4007-4021.
14. Kearney, J. F. & Lawton, A. R. (1975) *J. Immunol.* 115, 671-676.
15. Raff, M. D., Megson, M., Owen, J. J. T. & Cooper, M. D. (1976) *Nature (London)* 259, 224-226.
16. Wabl, M. R. & Burrows, P. D. (1984) *Proc. Natl. Acad. Sci. USA* 81, 2452-2455.

THIS PAGE BLANK (USPTO)

Functional immunoglobulin M production after transfection of cloned immunoglobulin heavy and light chain genes into lymphoid cells

(protoplast fusion/C418 selection)

ATSUO OCHI^{††}, ROBERT G. HAWLEY^{*†}, TERESA HAWLEY^{*†}, MARC J. SHULMANT[‡], ANDRÉ TRAUNECKER[§], GEORGES KÖHLER[§], AND NOBUMICHI HOZUMI^{*†}

^{*}Ontario Cancer Institute and [†]Department of Medical Biophysics, University of Toronto, 500 Sherbourne Street, Toronto, ON M4X 1K3 Canada; [‡]Rheumatic Disease Unit, Wellesley Hospital, Toronto, ON M4Y 1J3 Canada; and [§]Basel Institute for Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland

Communicated by Niels Kaj Jerne, July 11, 1983

ABSTRACT The rearranged immunoglobulin heavy (μ) and light (κ) chain genes cloned from the Sp6 hybridoma cell line producing immunoglobulin M specific for the hapten 2,4,6-trinitrophenyl were inserted into the transfer vector pSV2-neo and introduced into various plasmacytoma and hybridoma cell lines. The transfer of the μ and κ genes resulted in the production of pentameric, hapten-specific, functional IgM.

Work over the last decades has provided extensive information on immunoglobulin function and structure (1). Despite this information, it has been possible only in gross terms to relate molecular function with particular structural features.

With the advent of genetic engineering and gene transfer techniques, questions regarding structure-function relationships can now be readily addressed—that is, virtually any gene segment can be modified precisely *in vitro* and the novel segment can then be exchanged with its normal counterpart. By introducing such engineered genes into the appropriate cells, the effects of systematic alterations in protein structure on protein function can be assessed.

Because immunoglobulin production is a specialized function of cells of the B-lymphocyte lineage, it is expected that the conditions for proper Ig gene expression will be provided only in appropriate immunocompetent cells. For example, to produce normal pentameric IgM(κ), a cell must transcribe, process, and translate RNA for the μ and κ chains and also provide J protein, enzymes for the proper polymerization and glycosylation of the Ig chains, as well as a suitable secretory apparatus. We have previously described a system for transferring a functional immunoglobulin κ light chain gene into IgM-producing hybridoma cells (2). Here we extend this work to show that the transfer of the μ and κ chain genes of a defined specificity into various plasmacytoma and hybridoma cell lines results in the production of functional pentameric, hapten-specific IgM(κ).

MATERIALS AND METHODS

Cell Lines. X63Ag8 was originally derived (3) from the plasmacytoma MOPC21 and synthesizes IgG1(κ) of unknown specificity. X63Ag8.653 was derived from X63Ag8 as a subclone that synthesizes neither the heavy (γ 1) nor light (κ) chain (4). Similarly, Sp2/0Ag14 is an Ig nonproducing subclone of the Sp2 hybridoma (5). Sp6 is a hybridoma making IgM(κ) specific for the hapten 2,4,6-trinitrophenyl (TNP); originally this cell line produced the γ 1 and κ chains of X63Ag8 as well as the (TNP specific) μ_{TNP} and κ_{TNP} chains (6). A subclone of Sp6 not mak-

ing the γ 1 chain was isolated, and the Sp602 and Sp603 cell lines were derived from this γ 1 nonproducer. The mutant cell line igm-10, derived from Sp602 (7), lacks the gene encoding μ_{TNP} (8).

Gene Transfer. The construction of pSV2-neo plasmid vectors carrying the genes for μ_{TNP} or κ_{TNP} or both is described in the text. The vectors were transfected into the $r_k^- m_k^-$ *Escherichia coli* strain K803. To transfer the vector, bacteria bearing the appropriate plasmids were converted to protoplasts and fused to the indicated cell lines as described (2). The frequency of G418-resistant transformants per input cell was approximately 10^{-4} for X63Ag8 and Sp2/0Ag14, 10^{-5} for igm-10, and 10^{-6} for X63Ag8.653.

Analysis of Ig. As described previously (7), Ig was biosynthetically labeled, in the presence or absence of tunicamycin, immunoprecipitated, and analyzed by NaDodSO₄/polyacrylamide gel electrophoresis with or without disulfide bond reduction. TNP binding IgM was assayed by TNP-dependent hemagglutination and by TNP-dependent enzyme-linked immunoadsorbent assay (ELISA) as described (2, 7). The hemolyses of protein A-coupled erythrocytes and TNP-coupled erythrocytes were used to assay total IgM- and TNP-specific complement activating IgM, respectively (7).

Analysis of RNA and DNA. Cytoplasmic RNA was isolated according to Schibler *et al.* (9) and subjected to RNA blot analysis as described by Thomas (10).

Procedures for DNA extraction (11), nitrocellulose blotting (12), and radiolabeling of probes (13) have been described (14, 15). Probes specific for genes encoding immunoglobulin constant and variable regions are detailed in the figure legends.

RESULTS

Description of Vectors and Expression Systems. The hybridoma cell line Sp6 secretes IgM(κ) specific for the hapten TNP. We have previously described the cloning of the TNP-specific κ gene, designated Tk1 (16), and the construction of the recombinant, pR-Tk1, where Tk1 is inserted in the *Bam*HI site of the vector pSV2-neo (2, 17). The μ_{TNP} gene was cloned in λ Ch4A from *Eco*RI partially digested DNA of Sp6 cells, and this clone is designated Sp6-718. The 16-kilobase-pair (kbp) fragment carrying the variable and constant regions was obtained from Sp6-718 after partial digestion with *Eco*RI and was inserted at the *Eco*RI site of the vectors pSV2-neo and pR-Tk1. In these recombinants, designated pR-Sp6 and pR-HL_{TNP}, re-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: TNP, 2,4,6-trinitrophenyl; ELISA, enzyme-linked immunoadsorbent assay; kbp, kilobase pair(s); SV40, simian virus 40; kb, kilobase(s).

THIS PAGE BLANK (USPTO)

spectively, the μ_{TNP} gene lies in the same orientation as the κ_{TNP} gene in pR-Tk1—i.e., the direction of transcription of μ_{TNP} is opposite that of the simian virus 40 (SV40) early promoter (Fig. 1).

The mutant cell lines igk-14 and igm-10 that lack the κ_{TNP} gene and μ_{TNP} gene, respectively, were originally isolated from subclones of Sp6 (7). We have previously used igk-14 as a recipient cell line to assay expression of the κ_{TNP} gene (2). Expression of the μ_{TNP} gene of pR-Sp6 was assayed here in igm-10. The simultaneous production of both μ_{TNP} and κ_{TNP} chains from the vector pR-HL $_{TNP}$ is assayed in X63Ag8, the IgG1-producing plasmacytoma parent of the Sp6 hybridoma. In later experiments the pR-HL $_{TNP}$ vector was assayed in the non-producing cell lines Sp2/OAg14 and X63Ag8.653. IgM production by the transformants is compared with Sp603, a subclone of the Sp6 hybridoma.

Selection of IgM(κ)-Positive Transformants. The recombinant plasmid vectors bearing the Ig genes also contain the bacterial gene *neo*, which renders the recipient cells resistant to

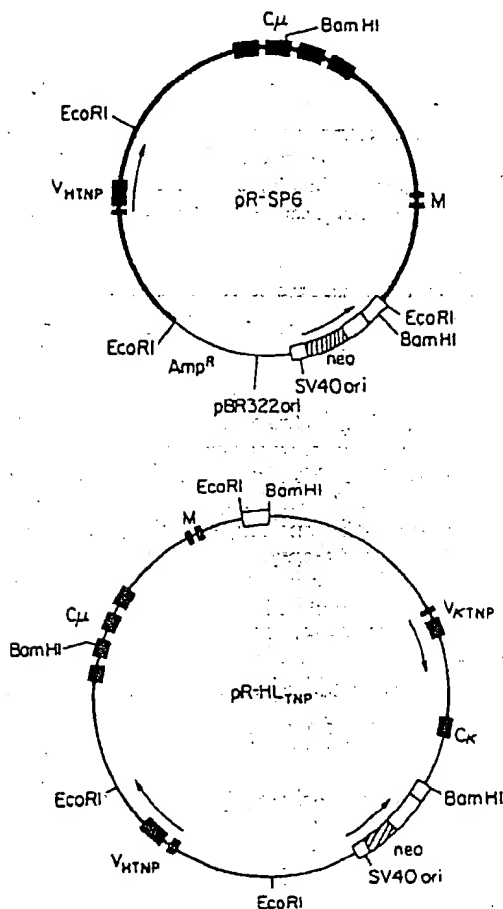


FIG. 1. Structure of the pR-Sp6 and pR-HL $_{TNP}$ plasmids. pR-Sp6 contains the functionally rearranged μ_{TNP} gene (≈ 16 kbp), which was inserted into the *EcoRI* site of pSV2-*neo* (see text). In addition to the μ_{TNP} gene, pR-HL $_{TNP}$ contains the functionally rearranged κ_{TNP} gene (9.6 kbp) at the *BamHI* site (2). Ig genes are represented by heavy dark lines. The directions of transcription of the Ig genes and the SV40 early region are indicated by arrows. The μ and κ exons are shown as filled boxes. M denotes alternative COOH-terminal coding regions that are utilized in the synthesis of membrane IgM. Thin lines are of pBR322 origin. The white boxes denote DNA derived from SV40, into which the bacterial gene conferring neomycin resistance (hatched box) has been inserted. For specific details concerning the pSV2-*neo* transfer vector (donated by P. Berg), see ref. 17.

the antibiotic G418 (17). To transfer the Ig genes into the hybridoma and plasmacytoma cells, bacteria harboring the recombinant plasmids were converted to protoplasts and fused with the various cell lines and G418-resistant cells were selected. Depending on the cell line, the efficiency of G418-resistant colonies ranged between 10^{-4} and 10^{-6} per input hybridoma or plasmacytoma cell (see *Materials and Methods*). The culture supernatant of G418-resistant colonies was tested for TNP-specific IgM by using either a TNP-specific ELISA or by assaying agglutination of TNP-coupled erythrocytes. In various experiments between 15% and 75% of the colonies were positive in such tests.

Analysis of μ_{TNP} and κ_{TNP} Production. Colonies that were positive for TNP-specific IgM were cloned by limiting dilution and examined further. The transformant IR44L1, derived from the κ_{TNP} -positive cell line igm-10 and the μ_{TNP} vector pR-Sp6, makes about 25% of the normal (Sp603) amount of IgM, as measured by the TNP-dependent ELISA. The transformant XR19L4, derived from the cell line X63Ag8 and the μ_{TNP} + κ_{TNP} vector pR-HL $_{TNP}$, makes about 10% of the normal amount of IgM.

To examine the μ_{TNP} and κ_{TNP} separately, these chains were radiolabeled and analyzed by NaDodSO₄/polyacrylamide gel electrophoresis (Fig. 2). The Sp603 hybridoma cell line still makes the κ chain of its plasmacytoma parent, X63Ag8 (Fig. 2, lane a), as well as the specific μ_{TNP} and κ_{TNP} chains (Fig. 2, lane e). The XR19L4 transformant derived from X63Ag8 has two additional bands (Fig. 2, lane b), which comigrate with the μ_{TNP} and κ_{TNP} of Sp603. The igm-10 cells used here make κ_{TNP} but have ceased to produce the κ of X63Ag8 (Fig. 2, lane c), presumably because of a rearrangement in this κ gene (see legend to Fig. 5). The IR44L1 transformant derived from igm-10 has one new band that comigrates with μ_{TNP} (Fig. 2, lane d). As shown in Fig. 3, analysis of unreduced IgM by NaDodSO₄/polyacrylamide gel electrophoresis indicates that the transformants make predominantly pentameric IgM [$(\mu_2\kappa_2)_2$].

RNA Production. To examine the RNAs expressed by the transferred μ_{TNP} and κ_{TNP} genes, cytoplasmic RNA from the transformants was fractionated by gel electrophoresis and probed

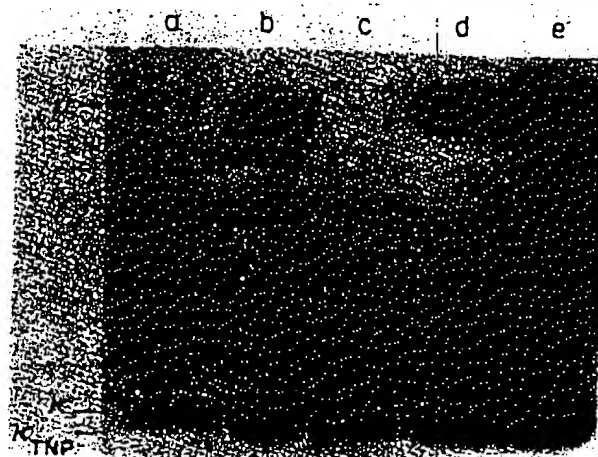


FIG. 2. Analysis of heavy and light chains of secreted Ig. G418-resistant transformant clones were biosynthetically radiolabeled with [¹⁴C]leucine as described (7). Secreted immunoglobulins were immunoprecipitated with rabbit anti-mouse IgM antibody complexed with protein A-Sepharose CL-4B beads (Pharmacia). The precipitated material was reduced with 2-mercaptoethanol and analyzed by electrophoresis on a NaDodSO₄/polyacrylamide gel. Lane a, X63Ag8; lane b, XR19L4; lane c, igm-10; lane d, IR44L1; and lane e, wild-type hybridoma Sp603.

THIS PAGE BLANK (USPTO)

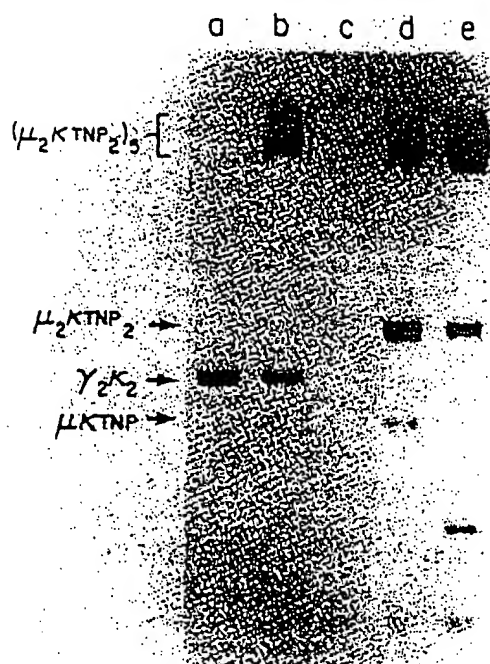


FIG. 3. Analysis of secreted (unreduced) Ig. The radiolabeled culture supernatants as described in the legend to Fig. 2 were analyzed by electrophoresis on a NaDodSO₄/polyacrylamide gel without reducing the disulfide bonds (7). Lane a, X63Ag8; lane b, XR19L4; lane c, igm-10; lane d, IR44L1; and lane e, wild-type hybridoma Sp603. The markers indicate the major forms of Sp603 IgM and X63Ag8 IgG1.

with various μ - and κ -specific DNA sequences (Fig. 4). RNA for the μ heavy chain was detected with a probe from the C_H4 region. The transformants XR19L4 and IR44L1 have bands at both 2.7 and 2.4 kilobases (kb), whereas the parental hybridoma Sp603 has only one band at 2.4 kb (Fig. 4A). A genomic probe containing the μ membrane-specific exon hybridized only to the 2.7-kb band (data not shown). RNAs of 2.7 and 2.4 kb have been found to encode the membrane (μ_m) and secreted (μ_s) forms of the μ chain, respectively (19–21). These results suggest that, whereas Sp603 makes RNA only for the μ_s form, the transformants make RNAs for both μ_m and μ_s . However, we have been unable to detect membrane IgM by staining with fluorescent μ -specific antibodies. The μ_m form has a longer polypeptide chain than does the μ_s form and consequently can be distinguished from μ_s by its lower mobility in NaDodSO₄/polyacrylamide gel electrophoresis. Therefore, we examined intracellular μ chains that were biosynthetically radiolabeled in the presence of tunicamycin; for each transformant we found only one μ band, and this band comigrated with the μ band of Sp6 (results not shown). These observations suggest that either the 2.7-kb RNA is not translated or that the μ_m protein is very short-lived in the transformants.

In a similar manner, the RNA blots were hybridized with a probe derived from the κ TNP V region. Compared to Sp603 and igm-10, the transformant XR19L4 was found to make a low amount of a 1.2-kb RNA that comigrated with authentic κ TNP RNA (Fig. 4B).

Structure of Transferred DNA. To analyze the organization of the transferred pR-Sp6 and pR-HL-TNP plasmids in the transformed cell lines, BamHI-digested cell DNA was hybridized with probes specific for the μ - and κ -chain constant region gene segments. The C_H1–2 probe used here spans the BamHI restriction site in the C_H2 exon (Fig. 1). Therefore, a minimum of two fragments is expected to be detected with this probe.

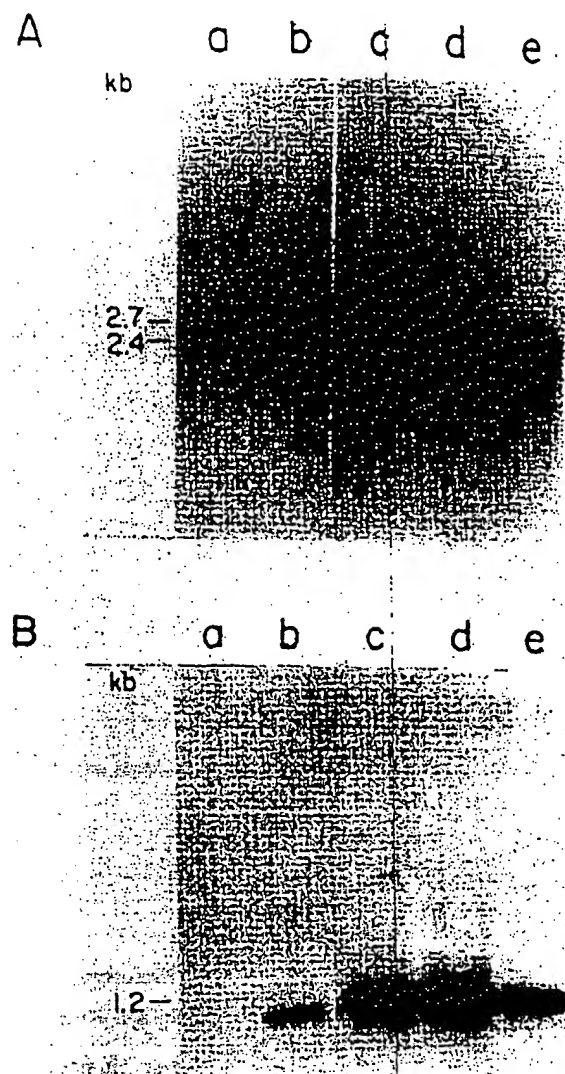


FIG. 4. Detection of μ TNP and κ TNP gene sequences in cytoplasmic RNA from transformed cell lines. Lanes a, X63Ag8; lanes b, XR19L4; lanes c, igm-10; lanes d, IR44L1; and lanes e, Sp603. Ten micrograms of total cytoplasmic RNA (9) was denatured with glyoxal, electrophoresed through a horizontal 1% agarose gel in 10 mM sodium phosphate buffer at pH 6.9, and transferred to nitrocellulose as described by Thomas (10). (A) The blot was hybridized with a ³²P-labeled probe corresponding to the C_H4 exon. This probe was isolated from the cDNA clone pH76 μ 17 (donated by J. Adams) after digestion with Pst I (18). (B) A similar blot was hybridized with a ³²P-labeled probe containing κ TNP V-region coding sequences (16). Sizes were estimated by comparison to mouse ribosomal 28S and 18S RNA (4.7 and 2.0 kb, respectively).

Two fragments of 6.0 and 16 kbp were detected in the DNA of both of the transformants. These correspond to the fragments generated by BamHI digestion of the intact pR-Sp6 and pR-HL-TNP plasmids (Fig. 5). In addition, one (XR19L4) or two (IR44L1) extra fragments could be detected in the DNA from these cell lines. In parallel experiments, sequences indicative of unintegrated pR-Tx1 plasmids have not been detected in the low molecular weight fraction of the Hirt supernatants (25) of similarly transformed igk-14 cells (results not shown). Taken together, these results suggest that the transferred genes are tandemly integrated into the chromosomal DNA of the recipient cells.

THIS PAGE BLANK (USPTO)

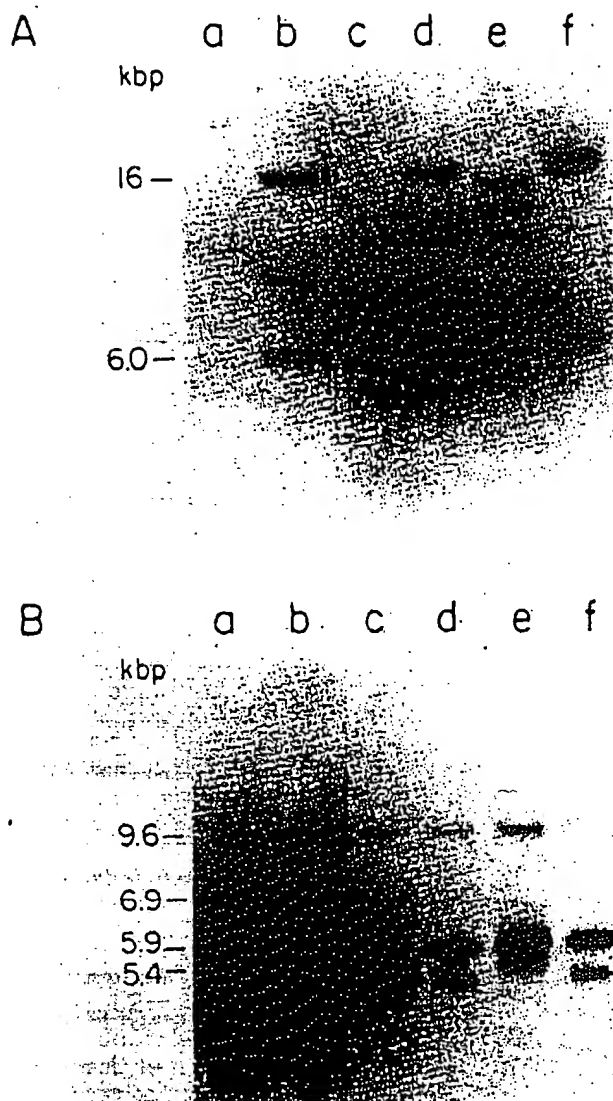


FIG. 5. (A) Detection of pR-Sp6 and pR-HL-TNP sequences in DNA from transformed cell lines. Lanes a, X63Ag8; lanes b, XR19L4; lanes c, igm-10; lanes d, IR44L1; lanes e, Sp603; and lanes f, igm-10 with 5 equivalents of pR-Sp6. *Bam*HI-digested DNA samples (20 μ g) were electrophoresed through a 1% agarose gel at 2 V cm^{-1} for 40 hr and transferred to nitrocellulose. (A) A previously hybridized blot (see B) was washed according to Thomas (10) and rehybridized to a ^{32}P -labeled probe containing the C₁ and C₂ exons. This probe was prepared by isolation of an appropriate fragment from a *Xba*I/*Hind*III digestion of a genomic clone of the μ -chain constant region gene segment. The bands corresponding to the μ -chain gene-containing fragments generated by *Bam*HI digestion of pR-Sp6 and pR-HL-TNP are indicated. The two bands observed in lane c (11 and 14 kbp) correspond to the functionally rearranged μ -TNP gene in the wild-type Sp603 cell line. (B) The same blot was hybridized with a ^{32}P -labeled probe containing the κ -constant region gene segment that was isolated from the plasmid pL21-5 (donated by R. Wall) (22). The bands at 9.6 kbp correspond to the κ -TNP gene (16). The bands at 6.9, 5.9, and 5.4 kbp correspond to rearranged κ chain genes present in the DNA of the X63Ag8 cell line (23, 24), two of which (5.9 and 5.4 kbp) were retained in the generation of the original Sp6 hybridoma. The 5.4-kbp band corresponds to the functionally rearranged X63Ag8 κ gene and this band is not observed in the case of igm-10 (lane c). Sizes were estimated by comparison to *Hind*III-digested λ phage DNA.

The pattern obtained for XR19L4 upon hybridization of the same blot with the C₂ probe is consistent with the above interpretation. DNA from this transformant contains a 9.6-kbp fragment corresponding to the wild-type κ -TNP gene (16) in addition

Table 1: Assay of functional IgM

Cell line	Phenotype	Hemolysis titer on erythrocytes coupled with		TNP/protein A ratio
		Protein A	TNP	
Sp603	IgM, κ (TNP) + κ (X63)	2 ^a	2 ^d	4
igm-10	κ (TNP)	<1	<1	—
IR44L1		2 ^b	2 ^d	4
X63Ag8	IgG1, κ	<1	<1	—
XR19L4		2 ^b	<1	<1.8
Sp2/OAg14	No Ig	<1	<1	—
SR1.2		2 ^a	2 ^d	4
SR40.1		2	2 ^d	2
X63Ag8.653	No Ig	<1	<1	—
X653R1.1		2 ^a	2 ^d	4

As described in the text, the transformants IR44L1 and XR19L4 were derived by introducing the μ -TNP gene alone or the μ -TNP and κ -TNP genes together into the igm-10 and X63Ag8 cell lines. Similarly, the cell lines SR1.2, SR40.1, and X653R1.1 were generated by transferring the μ -TNP + κ -TNP vector pR-HL-TNP into Sp2/OAg14 and X63Ag8.653. The indicated cell lines were grown to approximately 10^6 cells per ml, and culture supernatants were assayed for IgM concentration (lysis titer on protein A-coupled erythrocytes) and TNP-specific hemolysis activity (lysis titer on TNP-coupled erythrocytes). Culture supernatants were diluted serially 1:2 to obtain the end-point dilution (titer) that still caused lysis. The ratio of the TNP and the protein A titer is a measure of the specific activity of the secreted IgM.

to other fragments that correspond to the κ chain genes endogenous to the recipient X63Ag8 cell line (23, 24).

Assay of IgM Function. We have tested the normal functioning of the IgM produced by the transformants by assaying its action in complement-dependent lysis of TNP-coupled erythrocytes (Table 1). The IgM concentration in the culture supernatants of the indicated cell lines was measured by the hemolysis of protein A-coupled erythrocytes in the presence of anti-IgM (7). These results indicate that IgM made by IR44L1 has normal activity with regard to TNP binding and complement activation. However, the transformant XR19L4 makes IgM that has an activity that is less than 1/30th of the normal activity in the TNP-dependent hemolysis assay. X63Ag8 still produces the myeloma κ chain, and this κ chain can be incorporated into IgM, thus reducing TNP-specific hemolysis activity (7). To avoid this problem of the nonspecific myeloma κ chain, the μ -TNP + κ -TNP vector pR-HL-TNP was transferred into the nonproducer cell lines Sp2/OAg14 (5) and X63Ag8.653 (4). The IgM produced by transformants of these cell lines has normal activity for TNP-specific hemolysis (Table 1).

DISCUSSION

We and others have previously reported the expression of Ig light chain genes in various cell types (2, 26–29). In this paper we have described the construction of plasmids that bear genes for TNP-specific immunoglobulin μ and κ chains. The expression of these genes was studied after the transfer of the plasmids into various cell lines derived from Ig-secreting plasmacytomas or hybridomas. The transfer of these plasmids into these cells is usually (see below) sufficient to cause the production of pentameric IgM(κ) that binds antigen (TNP) and activates complement—that is, these cell lines (X63Ag8, X63Ag8.653, igm-10, and Sp2/OAg14) provide all of the machinery necessary for IgM production except the structural genes for the μ and κ

THIS PAGE BLANK (USPTO)

chains. The capacity to provide this machinery is present despite the fact that these cell lines have been propagated for years without overt selection for this property.

We expect that this system will be very useful in determining the structural requirements for normal IgM production and function. To date, the use of genetics for this purpose has been limited to the analysis of naturally occurring mutants that interfere with normal IgM processing and activity (7, 30). Although such mutants are useful as a starting point, *in vitro* mutagenesis offers a more rapid and systematic method of obtaining altered IgM. Thus, it should be possible to identify the amino acids that are critical for complement activation or Fc receptor binding. Similarly, one can expect to define the features that are necessary for pentamer formation, glycosylation, and secretion.

As is the case with other gene transfer systems, we have found that the various transformants produce quite different amounts of μ and κ chain, ranging from undetectable to approximately normal levels. In general, a linear relationship does not exist between the copy number of the transferred sequences and the level of Ig gene expression. Studies with transfer vectors presumed to be replication incompetent indicate that the transferred sequences integrate into different sites in the host chromosomes, independent of the method of transfer (31–33). Therefore, the context of the transferred genes is different from normal and different in each recipient. It is not known whether it is the different chromosomal locales that are responsible for the variation in the expression of the transferred genes or whether these results reflect a high frequency of mutation associated with the introduction of exogenous DNA into mammalian cells (34; 35).

The transformants XR19L4 and IR44L1 produce, in addition to a 2.4-kb RNA that comigrates with authentic μ RNA, a 2.7-kb RNA that appears to include the μ_m exon. As we have been unable to detect a μ_m protein, it is possible that the 2.7-kb RNA is aberrant in some respect (36–39). In contrast to the heavy chain gene results, the transferred κ chain genes in XR19L4 and in several transformants derived from *igk*-14 and the κ_{TNP} vector pR-T κ 1 (ref. 2; unpublished data) produce a single species of RNA that comigrates with authentic κ_{TNP} RNA.

We expect that the variations in the expression of the transferred genes will not interfere with the usefulness of this system in producing altered IgM for functional analysis. Furthermore, we anticipate that modifications of this protocol will allow investigation of the mechanisms controlling Ig gene expression.

Note Added in Proof. Gillies *et al.* (40) and Neuberger (41) have recently reported the expression of cloned heavy chain genes in transformed lymphoid cells.

We thank Nusrat Govindji and Catherine Filkin for expert technical assistance. This work was supported by grants from the Medical Research Council, the National Cancer Institute, the Arthritis Society, the Allstate Foundation, and Hoffmann-La Roche Ltd. A.O. was supported by a Terry Fox Cancer Research Fellowship from the National Cancer Institute. R.C.H. was supported by a studentship from the Medical Research Council.

1. Davies, D. & Metzger, H. (1983) *Annu. Rev. Immunol.* 1, 87–117.
2. Ochi, A., Hawley, R. G., Shulman, M. J. & Hozumi, N. (1983) *Nature (London)* 302, 340–342.
3. Köhler, C. & Milstein, C. (1975) *Nature (London)* 256, 495–497.
4. Kearney, J., Radbruch, A., Liesegang, B. & Rajewsky, K. (1979) *J. Immunol.* 123, 1548–1550.

5. Shulman, M., Wilde, C. & Köhler, G. (1978) *Nature (London)* 276, 269–270.
6. Köhler, C. & Milstein, C. (1976) *Eur. J. Immunol.* 6, 511–519.
7. Köhler, G. & Shulman, M. (1980) *Eur. J. Immunol.* 10, 467–476.
8. Köhler, G., Potash, M. J., Lehrach, H. G. & Shulman, M. J. (1982) *EMBO J.* 1, 555–563.
9. Schibler, U., Marcu, K. B. & Perry, R. P. (1978) *Cell* 15, 1495–1509.
10. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* 77, 5201–5205.
11. Gross-Bellard, M., Dudet, P. & Chambon, P. (1973) *Eur. J. Biochem.* 36, 32–38.
12. Southern, E. M. (1975) *J. Mol. Biol.* 97, 503–517.
13. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* 113, 237–251.
14. Hozumi, N., Hawley, R. G. & Murialdo, H. (1981) *Gene* 13, 163–172.
15. Hozumi, N., Wu, C. E., Murialdo, H., Roberts, L., Vetter, D., Fife, W. L., Whiteley, M. & Sadowski, P. (1981) *Proc. Natl. Acad. Sci. USA* 78, 7019–7023.
16. Hawley, R. G., Shulman, M. J., Murialdo, H., Gibson, D. M. & Hozumi, N. (1982) *Proc. Natl. Acad. Sci. USA* 79, 7425–7429.
17. Southern, P. J. & Berg, P. (1982) *J. Mol. Appl. Genet.* 1, 327–341.
18. Gough, N. M., Kemp, D. J., Tyler, B. M., Adams, J. M. & Cory, S. (1980) *Proc. Natl. Acad. Sci. USA* 77, 554–558.
19. Alt, F. W., Bothwell, A. L. M., Knapp, M., Siden, E., Mather, E., Koshland, M. & Baltimore, D. (1980) *Cell* 20, 293–301.
20. Rogers, J., Early, P., Carter, C., Calame, K., Bond, M., Hood, L. & Wall, R. (1980) *Cell* 20, 303–312.
21. Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. (1980) *Cell* 20, 313–319.
22. Wall, R., Gilmore-Hebert, M., Higuchi, R., Komaromy, M., Paddock, C., Strommer, J. & Salser, W. (1978) *Nucleic Acids Res.* 5, 3113–3128.
23. Storb, U., Arp, B. & Wilson, R. (1980) *Nucleic Acids Res.* 8, 4681–4687.
24. Walfield, A. M., Storb, U., Selsing, E. & Zentgraf, H. (1980) *Nucleic Acids Res.* 8, 4689–4707.
25. Hirt, D. (1967) *J. Mol. Biol.* 26, 365–369.
26. Rice, D. & Baltimore, D. (1982) *Proc. Natl. Acad. Sci. USA* 79, 7862–7865.
27. Oi, V. T., Morrison, S. L., Herzenberg, L. A. & Berg, P. (1983) *Proc. Natl. Acad. Sci. USA* 80, 825–829.
28. Falkner, F. C. & Zachau, H. G. (1982) *Nature (London)* 298, 286–288.
29. Picard, D. & Schaffner, W. (1983) *Proc. Natl. Acad. Sci. USA* 80, 417–421.
30. Shulman, M. J., Heusser, C., Filkin, C. & Köhler, G. (1982) *Mol. Cell. Biol.* 2, 1033–1044.
31. Robins, D. M., Ripley, S., Henderson, A. S. & Axel, R. (1981) *Cell* 23, 29–39.
32. de Saint Vincent, B. R., Delbruk, S., Eckhart, W., Meinkoth, J., Vitto, L. & Wahl, G. (1981) *Cell* 27, 267–277.
33. Folger, K. R., Wong, E. A., Wahl, G. & Capetchi, M. R. (1982) *Mol. Cell. Biol.* 2, 1372–1387.
34. Razzaque, A., Mizusawa, H. & Seidman, M. M. (1983) *Proc. Natl. Acad. Sci. USA* 80, 3010–3014.
35. Calos, M. P., Lebkowski, J. S. & Botchan, M. R. (1983) *Proc. Natl. Acad. Sci. USA* 80, 3015–3019.
36. Kemp, D. J., Harris, A. W. & Adams, J. M. (1980) *Proc. Natl. Acad. Sci. USA* 77, 7400–7404.
37. Alt, F. W., Rosenberg, N., Enea, V., Siden, E. & Baltimore, D. (1982) *Mol. Cell. Biol.* 2, 386–400.
38. Clarke, C., Berenson, J., Goverman, J., Boyer, P. D., Crews, S., Siu, C. & Calame, K. (1982) *Nucleic Acids Res.* 10, 7731–7749.
39. Nelson, K. J., Haimovich, J. & Perry, R. P. (1983) *Mol. Cell. Biol.* 3, 1317–1332.
40. Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. (1983) *Cell* 33, 717–728.
41. Neuberger, M. S. (1983) *EMBO J.* 2, 1873–1378.

THIS PAGE BLANK (USPTO)

Assembly of functional antibodies from immunoglobulin heavy and light chains synthesised in *E. coli*

Michael A. Boss, John H. Kenten, Clive R. Wood and J. Spencer Emtage

Celltech Ltd, 244 - 250 Bath Road, Slough SL1 4DY, Berks, UK

Received 5 March 1984; Revised and Accepted 16 April 1984

ABSTRACT

Genes for a murine μ heavy chain and a λ light chain immunoglobulin have been inserted into bacterial expression plasmids containing the *Escherichia coli* *trp* promoter and ribosome binding site. Induction of transcription from the *trp* promoter results in accumulation of both light and heavy chain polypeptides in appropriate host strains. Both proteins were found as insoluble products. Following extraction and purification of the immunoglobulin containing fractions, antigen binding activity was recovered. The activity demonstrates essentially the same properties as the antibody from the hybridoma from which the genes were cloned.

INTRODUCTION

Immunoglobulin genes and their products represent one of the most extensively studied families of eukaryotic macromolecules. Immunoglobulin polypeptides are secreted proteins and are synthesised with an amino-terminal signal peptide which is cleaved to yield the mature protein. The expression of immunoglobulin genes in *E. coli* forms the initial stage in the production of antibodies produced via recombinant DNA techniques. Such antibodies would have many uses. For example, detailed studies on antigen-antibody interactions following alterations of the antigen combining site by site directed mutagenesis could be carried out, or the Fc regions of the molecules could be altered for specific uses such as binding to matrices for immunopurification. Thus, it is surprising that with the many studies on expression of eukaryotic genes in *E. coli* (1), little has been done on immunoglobulin genes. So far immunoglobulin genes have been expressed in modified forms at low levels in *E. coli*, usually as incomplete amino-terminal fusion proteins (2,3,4). In one case, a *trpE*-IgE fusion has been expressed at 10% total *E. coli* protein (5).

Here we describe the bacterial expression of a murine μ heavy chain and a murine λ light chain immunoglobulin cDNA. The Ig μ and λ genes used in these studies are from cDNA clones isolated from the hybridomas B1-8 and S43

THIS PAGE BLANK (USPTO)

(6), respectively. These hybridomas were raised to the hapten 4-hydroxy-3-nitrophenyl acetyl (NP) and produce IgM antibodies which are termed heteroclitic, that is binding a related hapten e.g. 4-hydroxy-5-iodo-3-nitrophenyl acetyl (NIP) more strongly than NP (7). The μ gene was cloned from the hybridoma line B1-8 (8) and the λ gene was cloned from the related hybridoma S43 (9). However, the sequence of the λ from S43 varied by only two amino acids from B1-8 λ sequence, assuming that B1-8 λ has germ-line sequence (9). Both changes were conservative and outside of the complementarity determining regions. So in effect, the antibody from B1-8 can be used to represent the parental monoclonal antibody.

The two polypeptides were synthesised in *E. coli* as native proteins lacking eukaryotic signal sequences and presumably possessing amino-terminal methionine residues. High levels of expression were achieved using the *E. coli* K12 strain E103S or *E. coli* B but only low levels of expression occurred in HB101 (10). Following solubilisation of μ and λ polypeptides expressed in the same cell or different cells, the protein products were purified and antigen binding activity recovered. This activity demonstrates essentially the same properties as those found for an NP binding IgM hybridoma antibody.

MATERIALS AND METHODS

Chemicals and Cloning Procedures

Materials were purchased as follows: restriction enzymes (Bethesda Research Laboratories and New England Biolabs), T4 DNA polymerase (P-L Biochemicals), DNase I (Sigma), radioisotopes (Amersham), rabbit anti-mouse IgM (Bionetics), rabbit anti-IgM (Tago), rabbit anti- λ (Miles), MOPC 104E an IgM ($\mu\lambda_1$) myeloma protein (Bionetics), calf intestinal alkaline phosphatase and S1 nuclease (Boehringer Mannheim). Unless otherwise stated cloning procedures were as described (10).

Oligodeoxyribonucleotides were synthesised by the phosphotriester procedures (11) and were designed to have the sequences; 5'-GATCAATGCAGGCTGTTGTG-3' (R45), and 5'-ATTCCTGAGTCACAACAGCC-3' (R44).

Bacterial strains and Plasmids

Plasmids were transformed into *E. coli* strain HB101, DH1, *E. coli* B (10) and *E. coli* K12 strain E103S (Dr. Lee Simon, Waksman Institute of Microbiology, Piscataway, New Jersey 08854-0759, personal communication), and grown in L-broth containing 0.1g carbenicillin per litre. Plasmids pAB μ -11 (8) and pAB λ -15 (9) were a gift from Drs. A. Bothwell and D.

Baltimore.

Kari.

Pulse Chase

For pu
except that
Difco methi
CaCl₂ (22mg
exponential
methionine
added and t
Other Methc

Proced
(12), as we
Protein Pur

For fu
dissolved
dithiothre
(Pharmacia
in 9M urea
column was
eluted pea
peak of pr
Tris-HCl p
(Pharmacia
2.25-9M ur
bicarbonat

The μ
anion exch
Reconstitu

Produ
and light
supernatar
by sonicat
glycine-Na
dialysed f
glycine-Na
and 0.1mM
at 30,000

THIS PAGE BLANK (USPTO)

Baltimore. B1-8 proteins were gifts from Drs. M. Neuberger and T. Imanishi-Kari.

Pulse Chase Analysis

For pulse chase analysis inductions were set up as described above, except that the medium used consisted of: proline (0.3g/L), leucine (0.1g/L), Difco methionine assay medium (5g/L), glucose (60mg/L), thiamine (10mg/L), CaCl_2 (22mg/L), MgSO_4 (0.25g/L) and carbenicillin (0.1g/L). During exponential growth cells were pulse labelled with 30 $\mu\text{Ci/ml}$ L-[^{35}S] methionine for 2 minutes, after which unlabelled methionine (100 $\mu\text{g/ml}$) was added and the incubation continued for the times indicated.

Other Methods

Procedures used for bacterial lysis and fractionation were as described (12), as were procedures for inductions and protein assays (13).

Protein Purification

For further purification of the λ light chain, the cell debris were dissolved in 10mM Tris-HCl pH8.0, 25% formamide, 7M urea, 1mM EDTA and 2mM dithiothreitol. This material was loaded onto a DEAE Sephacel column (Pharmacia) (1 x 25cm at a flow rate of 5ml/hr) which had been equilibrated in 9M urea, 10mM Tris-HCl pH8.0, 1mM EDTA and 2mM DTT. The DEAE Sephacel column was developed using a 0-150mM NaCl gradient in loading buffer. The eluted peak of λ light chain immunoreactivity, corresponding to the major peak of protein was diluted to a final concentration of 2.25M urea, 10mM Tris-HCl pH8.0, 1mM EDTA, 2mM DTT and loaded onto an octyl-Sepharose column (Pharmacia) (2.5 x 10cm). Material was eluted by use of a urea gradient of 2.25-9M urea. The peak material was pooled, dialysed into ammonium bicarbonate and lyophilised.

The μ heavy chain was purified from 9M urea solubilised pellets by anion exchange chromatography and chromatofocussing (Pharmacia).

Reconstitution of Activity

Production of functional antibodies from E. coli expressing both heavy and light chains was achieved by lysing the cells and clarifying the supernatant by centrifugation. The insoluble material was washed, followed by sonication (3 times for 3 minutes), and finally dissolved in 9M urea, 50mM glycine-NaOH pH10.8, 1mM EDTA, and 20mM 2-mercaptoethanol. This extract was dialysed for 40 hours against 3 changes of 20 vols. of 100mM KCl, 50mM glycine-NaOH pH10.8, 5% glycerol, 0.05mM EDTA, 0.5mM reduced glutathione and 0.1mM oxidised glutathione. The dialysate was cleared by centrifugation at 30,000g for 15 minutes and loaded directly onto DEAE Sephacel, followed by

THIS PAGE BLANK (USPTO)

development with a 0-0.5M KCl linear gradient in 10mM Tris-HCl, 0.5mM EDTA, pH8.0.

The purified Ig μ and λ were treated as above, except that no anion exchange chromatography was carried out. The preparation was finally dialysed into phosphate buffered saline, 5% glycerol, 0.01% sodium azide and 0.5mM EDTA pH7.4.

RESULTS

Construction of Plasmids for Expression of λ Light Chain

We chose to express the λ gene in *E. coli* by direct expression of the gene lacking the eukaryotic signal peptide but containing a methionine initiator residue at the amino-terminus (met lambda). The approach used for bacterial synthesis of met- λ was to reconstruct the gene *in vitro* from restriction fragments of a cDNA clone and to utilise synthetic DNA fragments for insertion into the bacterial plasmid pCT54 (12) (figure 1). As a source of light chain we used the plasmid pAB λ 1-15 which contains a full-length λ_1 light chain cDNA cloned into the PstI site of pBR322 (9). We have previously outlined the construction of plasmids for the expression of λ light chain (14).

A plasmid was isolated (designated pCT54 19-1) and shown to have the anticipated sequence except that there was an alteration at the fifth codon from GTG to ATG, changing the amino acid at this point from valine to methionine (figure 1). Valine is an invariant residue at this position in mouse λ chains. Methionine, however, is the residue most frequently found in mouse κ chains at this position (15).

As most *E. coli* mRNAs have 6-11 nucleotides between the Shine-Dalgarno (SD) sequence and the AUG (16) the distance in pCT54 19-1 was reduced by modification at the ClaI site. Altering the distance between the SD sequence and the ATG has been demonstrated to alter the expression of a number of genes (13,14,17-20) presumably by placing the SD and ATG sequences in the optimal configuration for formation of the initiation complex. pCT54 19-1 was cut with ClaI and incubated with S1 nuclease. The amount of S1 nuclease was adjusted so that some DNA molecules would lose 1-2 extra base pairs as a result of 'nibbling' by the enzyme. This DNA on religation with T4 DNA ligase and transformation into *E. coli* strain HB101 gave rise to a number of plasmids which had lost the ClaI site. The nucleotide sequence across the modified region of two of these plasmids was determined (figure 1). pNP4 and pNP3 were shorter than pCT54 19-1 by 5 and 4 nucleotides

Figure 1.
chain in
Plasm
pAT153.
electroph
HindIII fr
fragments
yield plas
nuclease
E, EcoRI;

respective
structure
the SD or
either pC
have been
genes, no

THIS PAGE BLANK (USPTO)

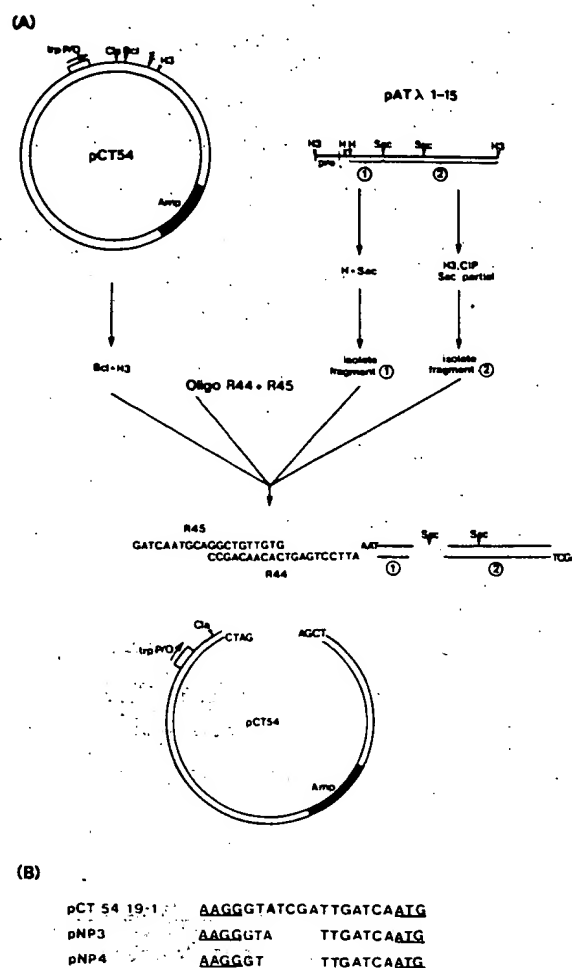


Figure 1. Construction of plasmids for the direct synthesis of λ light chain in *E. coli*.

Plasmid pAT λ 1-15 contains the λ gene inserted into the HindIII site of pAT153. (A) 5' HinfI - SacI fragment 1 was isolated by polyacrylamide gel electrophoresis. The 3' fragment 2 of the gene was isolated as a SacI - HindIII fragment. pCT54 was cut with BclI + HindIII and the λ gene fragments together with oligodeoxyribonucleotides R45 and R44 ligated to yield plasmid pCT54 19-1. (B) Digestion of pCT54 19-1 with ClaI and S1 nuclease produced plasmids pNP3 and pNP4, with reduced SD-ATG distances, E, EcoRI; H, HinfI; H3, HindIII.

respectively, giving SD-ATG distances of 9 and 10 nucleotides. Secondary structure analysis, as described (13), revealed no hairpin loop sequestering the SD or initiation codon into double-stranded regions of the mRNA of either pCT54 19-1 or the S1 derivatives. Such base pairing interactions have been shown drastically to reduce translational efficiency of a number of genes, notably those for phage λ *cro* (17), fibroblast and leukocyte

THIS PAGE BLANK (USPTO)

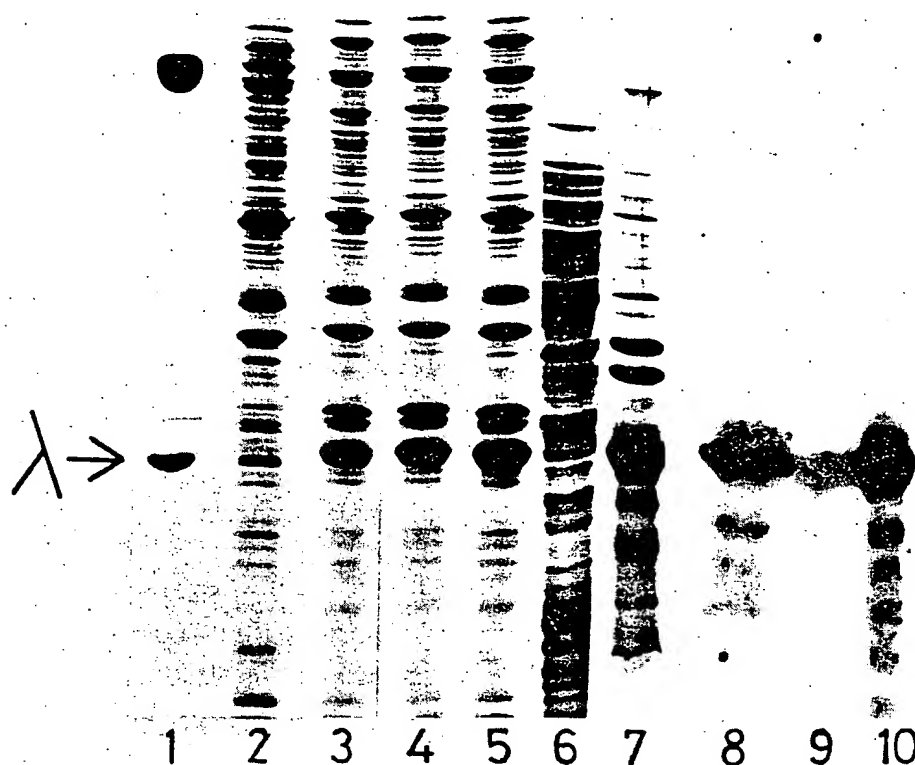


Figure 2. Accumulation and distribution of λ protein from *E. coli* E103S.

E103S cells containing pNP3 were grown under inducing conditions and samples taken throughout the induction cycle. Following electrophoresis, gels were stained with Coomassie blue (lanes 1-7) or subjected to analysis by Western blotting (lanes 8-10) using rabbit anti- λ serum and iodinated protein A ($2\mu\text{Ci}/\text{ml}$). Lane 1, purified MOPC104E myeloma protein standard indicating the position of authentic λ protein; lane 2, E103S containing pNP3 after growth to stationary phase in L-broth; lanes 3-5, samples from E103S containing pNP3 taken at increasing absorbance during induction; lanes 6 and 7, respectively, soluble and insoluble fractions from pNP3 containing E103S; lane 8, unfractionated extract; lanes 9 and 10, respectively, soluble and insoluble fractions.

interferons (18), SV40 small-t antigen gene (19) and a murine μ heavy chain (13).

Expression of λ Protein

HB101 cells containing pNP3 were found to express λ light chain as determined by immunoprecipitation (14) and Western blot analyses. No such protein was evident in extracts from pCT54 19-1, cells containing pNP4 which had a SD-ATG distance one base pair shorter than that of pNP3. HB101 expressed only a very low level of λ light chain (14).

In the absence of specific immunoprecipitation a novel protein band was not visible from extracts of HB101 containing pNP3 nor was the λ protein

found to accumulate. The difference in protein was stationary. These cells were found in HB101 under light of foreign protein required. These proteins by scanning the gel showed that the represented protein was 20 minutes after induction in E103S, suggesting a stationary strain.

pNP3, λ light chain but not in Coomassie

band as λ

8-10). T

containing

expressing

detected (

small amount

The presence

than full-

proteolytic

transcript

Expression

The

full-length

described

were grown

prepared,

the gel w

insoluble

negative

THIS PAGE BLANK (USPTO)

found to accumulate (data not shown). However, there was a dramatic difference when pNP3 was induced in the K12 strain E103S. In this strain λ protein was found to accumulate during induction until the cells reached stationary phase (figure 2, lanes 3-5) to a level of about 150 times that found in HB101 as determined by an ELISA (enzyme linked immunosorbent assay). These cells were found to contain inclusion bodies which appeared refractile under light microscopy, a phenomenon characteristic of high level expression of foreign proteins (21). An estimate of the percentage of total *E. coli* protein represented by recombinant λ protein was obtained by separating the proteins by gel electrophoresis, staining them with Coomassie blue and scanning the stained gel with a Joyce-Loebl chromoscan 3. This method showed that λ was the major protein present (figure 2, lane 5) and represented 13% of total *E. coli* protein. The λ protein had a half-life of 20 minutes in HB101 (data not shown) but accumulated to very high levels in E103S, suggesting that the lambda protein was much more stable in the latter strain. After cell lysis and centrifugation of HB101 or E103S containing pNP3, λ light chain was detected in the insoluble (figure 2, lanes 7 and 10) but not in the soluble fractions (figure 2, lanes 6), as determined by Coomassie blue staining. The identity of the major Coomassie blue stained band as λ protein was confirmed by Western blot analysis (figure 2, lanes 8-10). The presence of such immunoreactive bands was specific to pNP3 containing cells. When extracts from cells containing pCT70, a prochymosin expressing plasmid (12), were subjected to the same analysis, no bands were detected (data not shown). This more sensitive technique showed that a small amount of the λ protein was in the soluble fraction (figure 2, lane 9). The presence of a number of distinct immunoreactive proteins all smaller than full-length λ protein were also detected. These may result from proteolytic degradation of λ protein, from premature termination of transcription or from internal initiation of translation.

Expression of μ Protein

The construction of plasmids pNP11 and pNP14 for the expression of full-length mature μ protein under the control of the *trp* promoter has been described (13). *E. coli* B cells containing the μ expression plasmid pNP11 were grown under inducing conditions and soluble and insoluble extracts prepared, and analysed by SDS-PAGE. A novel band was seen after staining the gel with Coomassie blue in the lane containing proteins from the insoluble fraction (figure 3, lane 2). This band was not seen in the negative control lane which contained proteins from the same fraction from

THIS PAGE BLANK (USPTO)

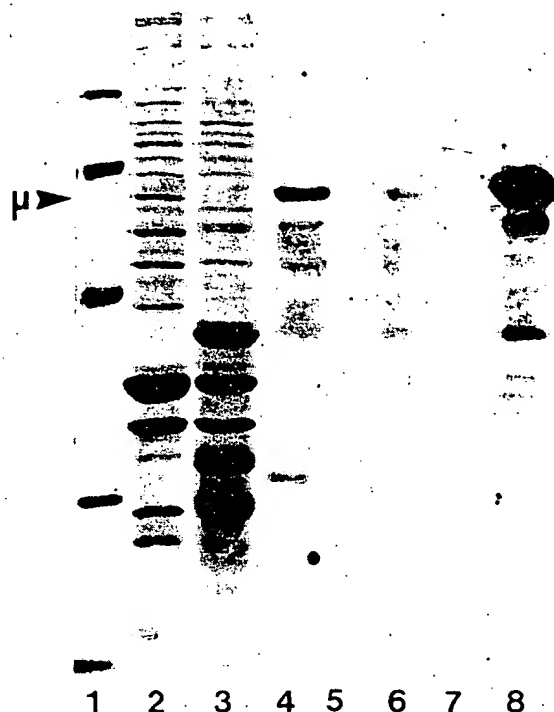


Figure 3. Expression and distribution of μ protein from *E. coli* B.

E. coli B cells containing pNP14 were grown under inducing conditions. Following electrophoresis, gels were stained with Coomassie blue (lanes 1-3) or subjected to analysis by Western blotting (lanes 4-8) using rabbit anti-IgM serum. Lane 1, molecular weight markers (94, 67, 43, 30 and 20Kd); lanes 2 and 4, respectively, pNP14 containing insoluble fraction; lanes 3 and 5, pCT70 containing insoluble fraction; lanes 6 and 7, respectively, pNP14 and pCT70 containing soluble fractions; lane 8, unfractionated sample from cells containing pNP14.

cells harbouring pCT70 (figure 3, lane 3). The novel band was found to migrate to a position corresponding to a protein of a molecular weight within less than 5% of the actual molecular weight of non-glycosylated μ of 62.5Kd. A duplicate set of lanes were transferred to nitrocellulose, and Western blotted. Alignment of the stained gel and the blot autoradiogram confirms that this novel band is antigenically related to IgM (figure 3, lanes 4 and 8). No band was found in extracts from cells containing pCT70 (figure 3, lanes 5 and 7). Only a low amount of μ was found in the soluble fraction (figure 3, lane 6).

A greatly increased level of expression of μ was found in *E. coli* B compared to HB101. Pulse chase analysis demonstrated that in *E. coli* B, a similar level of μ protein was detected after a 60 minute chase (figure 4, lane 3) as was seen after the initial labelling period (figure 4, lane 1). In HB101, however, very little μ protein could be seen after a 10

Figure 4.

E. coli B cells containing pNP14 were grown under inducing conditions. Following electrophoresis, gels were stained with Coomassie blue (lanes 1-3) or subjected to analysis by Western blotting (lanes 4-8) using rabbit anti-IgM serum. Lane 1, molecular weight markers (94, 67, 43, 30 and 20Kd); lanes 2 and 4, respectively, pNP14 containing insoluble fraction; lanes 3 and 5, pCT70 containing insoluble fraction; lanes 6 and 7, respectively, pNP14 and pCT70 containing soluble fractions; lane 8, unfractionated sample from cells containing pNP14.

minute chase

(figure 4, lane 3)

examined by

Purification

The purified

from the

For

in a Tris

DEAE Sepha

9M urea, 1

0-150mM Na

eluted pe

was diluted

THIS PAGE BLANK (USPTO)

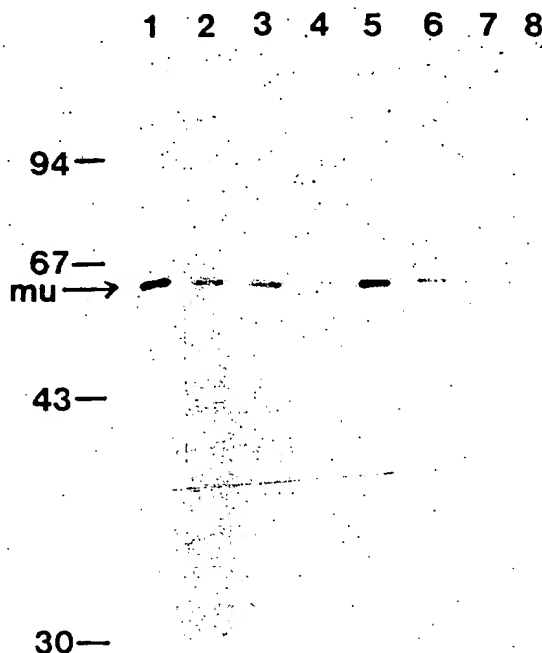


Figure 4. Pulse chase autoradiogram of pNP14 in *E. coli* B and HB101.

E. coli B and HB101 cells harbouring pNP14 were grown under inducing conditions and pulsed with L-[³⁵S]-methionine for 2 mins. after which unlabelled L-methionine was added to 100 µg/ml (zero time). Cells were harvested at varying times, and samples analysed by SDS-PAGE following immunoprecipitation. *E. coli* B: lane 1, zero time; lane 2, 30 min; lane 3, 60 min. Lane 4 is a zero time sample for pCT70 in HB101. HB101: lane 5, zero time; lane 6, 5 min; lane 7, 10 min; lane 8, 30 min.

minute chase (figure 4, lane 7), and none after 30 minutes (figure 4, lane 8), compared to the amount detected after the initial labelling period (figure 4, lane 5). Induced *E. coli* B cells harbouring pNP14 when examined by phase contrast microscopy were found to contain inclusion bodies.

Purification of Recombinant

The presence of λ light chain in the insoluble fraction was a useful purification step since it both concentrated the protein and separated it from the bulk of *E. coli* soluble proteins.

For further purification the *E. coli* insoluble material was solubilised in a Tris-HCl buffer containing 25% formamide and 7M urea and loaded onto a DEAE Sephacel column which had been preequilibrated in the same buffer with 9M urea, but without formamide. The bound material was eluted using a 0-150mM NaCl gradient in the urea buffer. The λ protein was the major eluted peak as determined by gel electrophoresis and ELISA. The λ protein was diluted to a final concentration of 2.25 urea and loaded onto an octyl-

THIS PAGE BLANK (USPTO)

Sephacrose column. The λ protein bound and was eluted by using a 2.25-9M urea gradient. Following this step, only a single band of Coomassie blue stainable material corresponding to recombinant λ protein was visualised by SDS-PAGE (data not shown).

Expression of μ and λ Polypeptides in the same cell

Each of the Ig μ and λ genes in expression plasmids were transformed into the same *E. coli* cell to direct the synthesis of both Ig μ and λ polypeptides. In order to overcome plasmid incompatibility and provide a second antibiotic resistance marker, the *trp* promoter and λ sequences were excised from pNP3 on a HindIII-Bam HI fragment and inserted into the HindIII-Bam HI fragment of pACYC184 (10) to create pACYC λ . HB101 cells containing this plasmid were found to grow very poorly. This weak growth was thought to result from read through of RNA polymerase into the origin of replication, and inhibition of growth was virtually eliminated, by inserting the bacteriophage T7 early transcriptional terminator (22) at the HindIII site of pACYC λ . The resultant plasmid pAC λ T7-1 has a chloramphenicol resistance gene and an origin compatible with the pBR322-derived origin on pNP14, the Ig μ expressing plasmid. Transformation of both plasmids into the same *E. coli* B cell was achieved in two steps, firstly pNP14 was introduced, and then pAC λ T7-1, in two separate transformations to give ampicillin and chloramphenicol resistant clones.

E. coli B cells derived from double-transformant clones showed the presence of inclusion bodies and two novel polypeptide bands on stained gels of the insoluble fraction after lysis. These two bands correlated both with immunological activity by Western blotting for Ig μ and λ and their expected molecular weights of 63,500 and 25,000 daltons respectively (data not shown). Stability of the plasmids was also investigated and it was found that after 36 hours in shake flasks only 5% of the *E. coli* contained both antibiotic resistance markers, although 35% were carbenicillin resistance and 74% were chloramphenicol resistant. This illustrates the selective pressure against both plasmids together.

Reconstitution of Antigen Binding

It was of great interest to determine whether the concomitant expression of μ and λ would lead to the formation of functional IgM. In order to determine this extracts were made from *E. coli* containing both Ig μ and λ polypeptides and these tested for antigen binding. We used a two-site sandwich ELISA which detects μ chain binding to haptenated bovine serum albumin (NIP-caproate-BSA). This assay demonstrates

Figure 5. P
Fractio
expressed Ig
equivalents;
protein dete

sensitivity
and insolubl
same buffer
for assay.

In orde
the insolubl
disulphide
assays of ma
obtained.

any detaile
exchange ch
of signific
to BSA (fig
expressed a
was not fou
not shown).
The separa
length Ig
contained v

The p
E. coli pro
binding ac
property o
factor, or

THIS PAGE BLANK (USPTO)

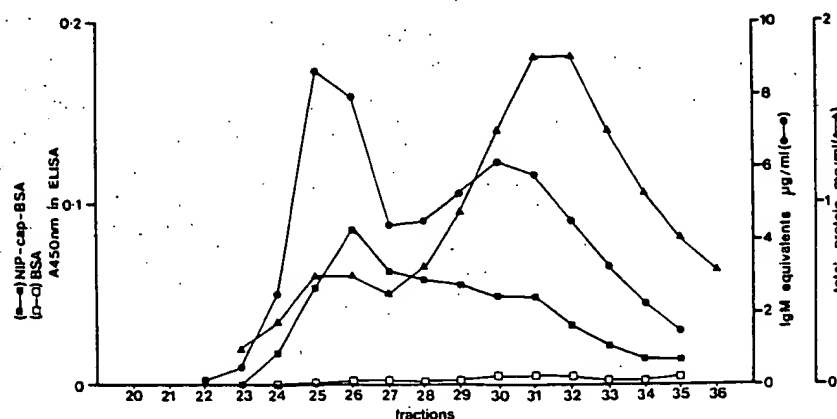


Figure 5. Purification of reconstituted antibodies.

Fractions from DEAE Sephacel anion exchange chromatography of *E. coli* B expressed Ig μ and λ . Analysis shows level of Ig μ expressed as B1-8 IgM equivalents; NIP-cap-BSA and BSA binding activities from ELISA's; total protein determined by A280nm readings (1.5 A280=mg/ml).

sensitivity to 60 pg of B1-8 IgM. The extracts were prepared as soluble and insoluble material. The insoluble material was solubilized in the same buffer used in lysis but containing 8M urea followed by its dilution for assay. However, no antigen binding activity was detected.

In order to obtain activity for the Ig μ and λ , extracts were made of the insoluble fraction and these dialysed into buffer conditions in which disulphide interchange will occur at a higher frequency. The results from assays of material processed in this way indicated that some activity was obtained. The level of activity obtained in this way was too low to do any detailed studies on, so the resultant dialysate was purified by anion exchange chromatography (figure 5). This process resulted in the isolation of significant NIP-cap-BSA binding activity over that of background binding to BSA (figure 5). The assay of the fractions for the level of Ig μ , expressed as B1-8 IgM equivalents demonstrated two peaks of activity. This was not found to correlate with full length Ig μ by Western blotting (data not shown). The first peak observed may represent a fragment of Ig μ . The separation of NIP-cap-BSA binding activity from the majority of full length Ig μ and protein indicates that the hapten binding activity is contained within a particular molecular species formed at low efficiency.

The processing of insoluble material obtained from Ig μ expression in *E. coli* produced a similar IgM protein profile but without NIP-cap-BSA binding activity. This demonstrates that the activity recovered was a property of the combined immunoglobulin expression, not of some *E. coli* factor, or of the Ig μ heavy chain alone.

THIS PAGE BLANK (USPTO)

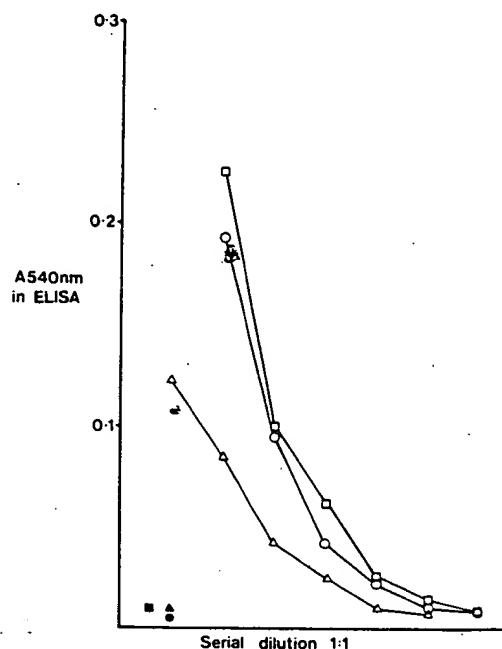


Figure 6. Specific hapten binding of reconstituted antibodies.

NIP-cap-BSA binding activity from fraction 26 (Δ), purified Ig μ and λ (\square) and B1-8 (\circ). Binding in the presence of the 30 μ M NIP-cap. (\blacktriangle , \blacksquare , \bullet respectively).

Further studies of the characteristics of the hapten (NIP-cap-BSA) binding were carried out. The reduction of hapten binding activity on dilution of renatured μ and λ expressed and purified together was less than that found for either B1-8 IgM or purified μ and λ expressed separately (figure 6). The dilution curves for B1-8 antibody and separately expressed μ and λ were virtually identical. Free hapten was found to inhibit most of the binding activity in both undiluted and diluted samples. Using B1-8 antibody as a standard for both IgM and hapten binding, the specific activity of the assembled antibody was calculated to be 1.4×10^4 gm/gm of IgM equivalents. This value demonstrates the inefficient recovery of activity, but possibly represents an underestimate of the specific activity due to an overestimate of full-length Ig μ in these fractions, as described above.

Heteroclitc Nature of Recombined Antibody

Detailed specificity of binding to NIP-cap-BSA was investigated by comparing the assembled antibodies with B1-8 IgM in the presence of free NIP-cap and NP-cap (figure 7). Both B1-8 IgM and the assembled antibodies showed that higher NP-cap than NIP-cap concentrations were required to inhibit NIP-cap-BSA binding.

Figure 7.
Bindi
purified I

The h
NP at 50%
(I50) were
(Table 1).
(Table 1).

The s
lower than
were both
concentrat
inhibit b:
Although
 λ are very
inhibition

Tabl
anti

B1-8
Fra
Fra
Pur
SD

THIS PAGE BLANK (USPTO)

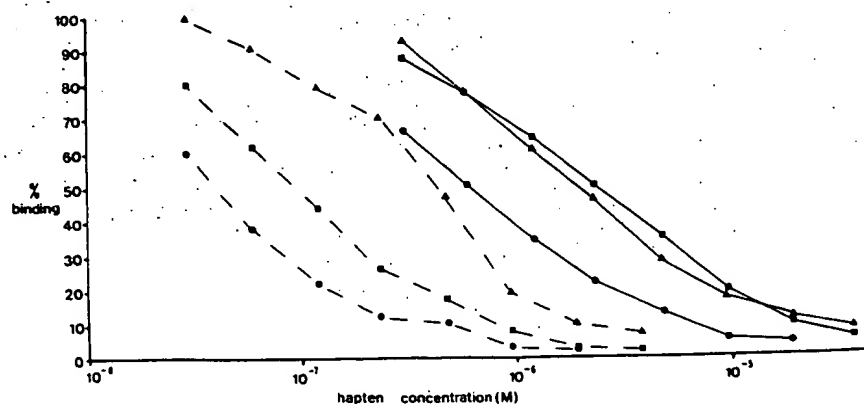


Figure 7. Heteroclitic binding of reconstituted antibody.

Binding of antibodies to NIP-cap-BSA; B1-8 IgM (\blacksquare), fraction 26 (\blacktriangle), purified Ig μ and λ (\bullet), in the presence of free NIP-cap (---) or NP-cap (—).

The heteroclitic nature is demonstrated by the molar ratio of NIP to NP at 50% inhibition. The concentrations of NIP and NP at 50% inhibition (I50) were found to be similar for both B1-8 and the assembled antibodies (Table 1). Also the specificity ratios (NP I50/NIP I50) were similar (Table 1).

The specificity ratios for the μ and λ expressed together were somewhat lower than those found for B1-8 IgM or separately expressed μ and λ which were both very similar (Table 1). This is due to the identical concentrations of NP but the greater concentrations of NIP required to inhibit binding of μ and λ expressed together compared with B1-8 IgM. Although the specificity ratios for B1-8 IgM and separately expressed μ and λ are very similar, lower concentrations of NIP and NP are required for inhibition of separately expressed μ and λ compared to B1-8 IgM.

Table 1. Hapten concentration at 50% inhibition (I50) of binding of antibodies to NIP-cap-BSA solid phase.

	NIP I50 μ M	NP I50 μ M	NP I50 NIP I50
B1-8 IgM	0.13 (SD, 0.05)	3.7 (SD, 2.9)	29
Fraction 26	0.34 (SD, 0.09)	1.9 (SD, 0.4)	6
Fraction 27 and 28	0.11 (SD, 0.02)	1.1 (SD, 0.3)	10
Purified μ and λ	0.04	0.84	22

SD = Standard Deviation.

) .)

THIS PAGE BLANK (USPTO)

DISCUSSION

Low level expression of λ and μ polypeptides was demonstrated in *E. coli* HB101. A greater level of expression of λ was found in strains E103S and *E. coli* B (J. Schoemaker, personal communication) and represented 13% total *E. coli* protein. A higher μ concentration was also found in *E. coli* B than HB101 cells containing pNP14 and was equivalent to 1% total *E. coli* protein. These differences in steady-state levels of the μ and λ proteins produced is most likely to result from different levels of protein stability. λ protein had a half life of 20 minutes in HB101 but was stable and accumulated in E103S. Similarly the pulse chase data shows that μ is more stable in *E. coli* B than HB101. This increased stability in *E. coli* B may be explained by the formation of inclusions, perhaps compartmentalizing μ away from proteases. *E. coli* B is known to be deficient in lon protease (23,24), so that the absence of this protease either acting upon μ , or being responsible for activating other proteases that act on μ , may be the major factor resulting in accumulation in *E. coli* B.

Both μ and λ proteins were expressed on compatible plasmids in the same cells. This expression of two different polypeptides necessary for the formation of a higher eukaryotic multi-subunit protein in *E. coli* represents the first of its kind. Thus it was of great interest to see if this concomitant expression of Ig μ and λ would lead to the formation of complete and functional IgM. No functional antibody was found following solubilisation of insoluble material. This indicated that the Ig μ and λ were not covalently interacting in the insoluble material in a specific way and that the inclusions represented non-active protein. The lack of activity correlated with insignificant amounts of Ig μ or λ in the soluble fraction. However, only a low percentage of cells were found to contain both plasmids after 36 hours induction. Thus the failure to find active (and presumably soluble) antibodies *in vivo* might simply be a reflection of the possibility that both polypeptides were not expressed at high levels in the same cells, and only a low percentage of cells contained both plasmids anyway.

Functional antibody activity, as defined by antigen binding, was obtained following dialysis of extracts in conditions in which disulphide interchange occurs. Significant antigen binding activity was recovered, which was abolished by competition with free hapten in a binding assay. Serial dilution of samples revealed a similar gradient in the reduction of binding compared to the monoclonal B1-8. The bacterially synthesised

antibody also NIP in a manner functional across polypeptide chains of different cell dependent on folding the μ antigen binding

Although was low, those very similar the efficiency number of disulphide heavy and light and so the inter-disulphide bonds immunoglobulin the inter-chain with high efficiency the formation bacterially to the appearance for accurate oxidised λ

These produced in It is now possible interaction site-directed

ACKNOWLEDGMENTS

We thank PAB1-15 and B1-8 proteins during the

REFERENCES

1. Harris Vol.

THIS PAGE BLANK (USPTO)

antibody also demonstrated higher affinity binding to the related hapten NIP in a manner similar to that of the B1-8 antibody. Recovery of functional activity was not dependent on coexpression of the μ and λ polypeptide chains since results from purified μ and λ expressed from different cells showed similar results. Functional activity was, however, dependent on association between heavy and light chain polypeptides, since folding the μ heavy chain in the absence of λ light chain produced no antigen binding molecules.

Although the overall efficiency of production of functional antibodies was low, those molecules which did assemble, demonstrated activity in a way very similar to the hybridoma synthesised molecules. It is unclear why the efficiency of assembly was low, but was possibly due to the large number of disulphide bridges needed to assemble an IgM molecule. The heavy and light chain polypeptides were completely unfolded before assembly and so the intra-domain disulphide bridges as well as the inter-chain disulphide bridges need to be made correctly. Classically, when immunoglobulin heavy and light chains are separated and reassembled, only the inter-chain disulphide bridges are reduced and these molecules refold with high efficiency (25,26). The conditions used in our experiments for the formation of disulphide bridges were probably adequate, since the bacterially synthesised λ light chain when folded under such conditions led to the appearance of a discrete lower molecular weight species, as expected, for accurately folded and oxidised λ light chain, which comigrated with oxidised λ protein from B1-8 (data not shown).

These results indicate that non-glycosylated immunoglobulin molecules produced in E. coli can be used for studies concerning antigen-binding. It is now possible to investigate the effects on antibody-antigen interactions following mutation of specific amino-acids of the antibody by site-directed mutagenesis and expression of the polypeptides in E. coli.

ACKNOWLEDGEMENTS

We thank Drs. A. Bothwell and D. Baltimore for giving us plasmids pAB λ 1-15 and pAB μ -11; Drs. M. Neuberger and T. Imanishi-Kari for gifts of B1-8 protein; and many colleagues at Celltech who have given us advice during the course of this work.

REFERENCES

1. Harris, T.J.R. (1983) in Genetic Engineering, Williamson, R. Ed., Vol. 4, pp.127-183, Academic Press, London.

THIS PAGE BLANK (USPTO)

2. Amster, O., Salomon, D., Zemel, O., Zamier, A., Zeelon, E.P., Zantor, F. and Schechter, I. (1980) Nucl. Acids Res. 8, 2055-2065.
3. Kemp, D.J. and Cowman, A.F. (1981) Proc. Natl. Acad. Sci. USA 78, 4520-4524.
4. Kurokawa, T., Seno, M., Sasada, R., Ono, Y., Onda, H., Igarashi, K., Kikuchi, M., Sugino, Y. and Honjo, T. (1983) Nucl. Acids Res. 11, 3077-3085.
5. Kenten, J.H., Helm, B.A., Cattini, P.A. and Gould, H.J. (1984) Proc. Natl. Acad. Sci. USA. In press.
6. Reth, M., Hammerling, G.J. and Rajewsky, K. (1978) Eur. J. Immunol. 8, 393-400.
7. Imanishi, T. and Makela, O. (1973) Eur. J. Immunol. 3, 323-330.
8. Bothwell, A.L.M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewsky, K. and Baltimore, D. (1981) Cell 24, 625-637.
9. Bothwell, A.L.M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewsky, K. and Baltimore, D. (1982) Nature 298, 380-382.
10. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning (a laboratory manual), Cold Spring Harbor Laboratory.
11. Patel, T.P., Millican, T.A., Bose, C.C., Titmas, R.C., Mock, G.A. and Eaton, M.A.W. (1982) Nucl. Acids Res. 10, 5606-5620.
12. Emtage, J.S., Angal, S., Doel, M.T., Harris, T.J.R., Jenkins, B., Lilley, G. and Lowe, P.A. (1983) Proc. Natl. Acad. Sci. USA 80, 3671-3675.
13. Wood, C.R., Boss, M.A., Patel, T.P. and Emtage, J.S. Nucl. Acids Res. (submitted).
14. Boss, M.A. and Emtage, S. (1983) Gene expression, UCLA. Symposia on Molecular and Cellular Biology, New Series, Hamer, D. and Rosenberg, M. Eds., Vol. 8, pp 513-522, Alan R. Liss Inc., New York.
15. Kabat, E.A., Wu, T.T., Reid-Miller, M., Perry, H. (1983) Sequences of proteins of immunological interest. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health.
16. Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B.S. and Stormo, G. (1981) Ann. Rev. Microbiol. 35, 365-403.
17. Iserentant, D. and Fiers, W. (1980) Gene 9, 1-12.
18. Shepard, H.M., Yelverton, E. and Goeddel, D.V. (1982) DNA 1, 125-131.
19. Gheysen, D., Iserentant, D., Derom, C. and Fiers, W. (1982) Gene 17, 55-63.
20. Yelverton, E., Norton, S., Obijeski, J.F. and Goeddel, D.V. (1983) Science 219, 614-620.
21. Williams, D.C., Van Frank, R.M., Muth, W.L. and Burnett, J.P. (1982) Science 215, 687-689.
22. Dunn, J.J. and Studier, F.W. (1980) Nucl. Acids Res. 8, 2119-2132.
23. Donch, J. and Greenberg, J. (1968) J. Bacteriol. 95, 1555-1559.
24. Donch, J., Chung, Y.S. and Greenberg, J. (1969) Genetics 61, 363-370.
25. Bridges, S.H. and Little, J.R. (1971) Biochemistry 10, 2525-2530.
26. Cathou, R.E. and Dorrington, K.J. (1975) in Subunits in Biological Systems, Timasheff, S.N. and Fasman, G.D. Eds., Part C. pp 91-224, Marcel Dekker, Inc., New York.

Molecular am

Ian D.Hickso

Department

Received 13

ABSTRACT

The
fusions to
low and
product h
leftward
protein,
was purif

INTRODUCT

Muta
genetic
Mutants
DNase (2-
of molecu
identific
and recC

Pur
in low
enzyme,
presente
amount c
and recC
attribu
genes. F
availab
possible
RecB an
vectors
leftwa

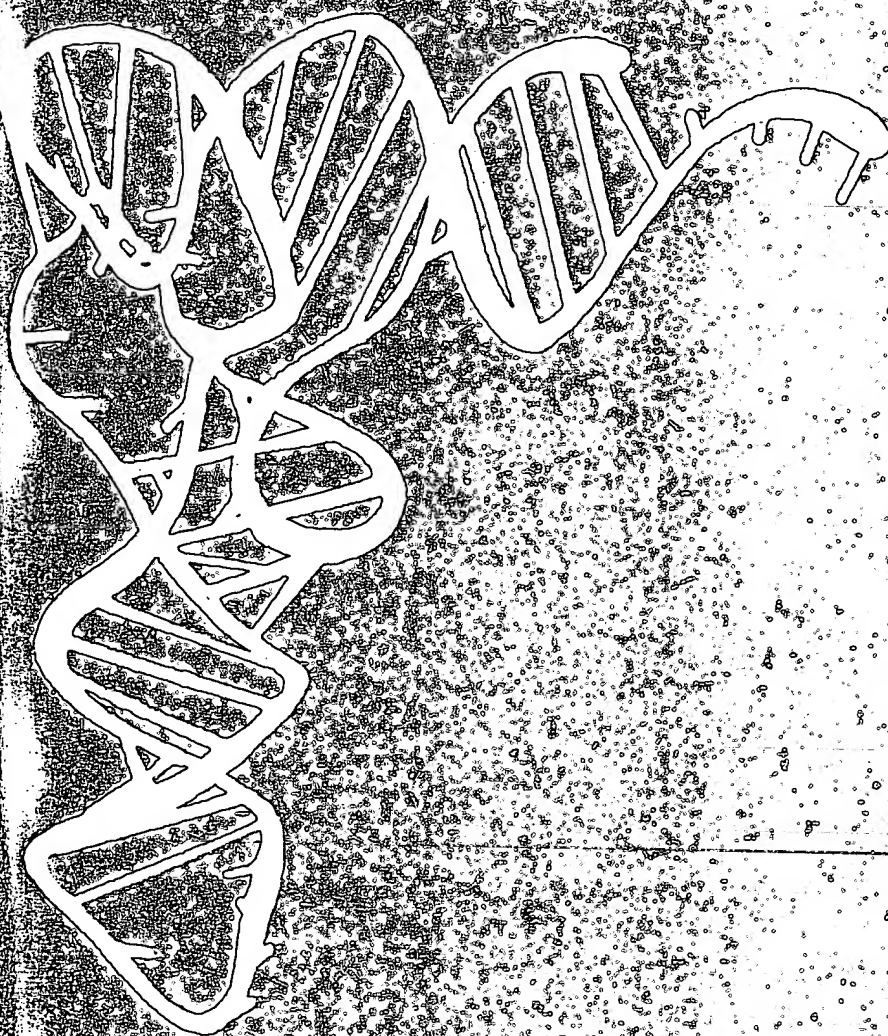
THIS PAGE BLANK (USPTO)

Nucleic Acids Research

STANFORD UNIVERSITY
LIBRARIES

JUN 5 1984

Volume 12 number 9 May 11 1984



IRL PRESS
Oxford · Washington DC

ISSN 0305 1048
Codon NARHAD

THIS PAGE BLANK (USPTO)

Secretion of functional antibody and Fab fragment from yeast cells

(*Saccharomyces cerevisiae*/chimeric antibody/yeast signal sequence/tumor antigen binding/effector functions)

ARNOLD H. HORWITZ*†, C. PAUL CHANG*, MARC BETTER*, KARL ERIK HELLSTROM†, AND RANDY R. ROBINSON*

*International Genetic Engineering, Inc., 1545 17th Street, Santa Monica, CA 90404; and †Oncogen, 3005 First Avenue, Seattle, WA 98121

Communicated by M. Frederick Hawthorne, August 11, 1988

ABSTRACT We have constructed yeast strains that secrete functional mouse-human chimeric antibody and its Fab fragment into the culture medium. For chimeric whole antibody, cDNA copies of the chimeric light-chain and heavy-chain genes of an anti-tumor antibody were inserted into vectors containing the yeast phosphoglycerate kinase promoter, invertase signal sequence, and phosphoglycerate kinase polyadenylation signal. Simultaneous expression of these genes in yeast resulted in secretion of properly folded and assembled chimeric antibody that bound to target cancer cells. Yeast chimeric antibody exhibited antibody-dependent cellular cytotoxicity activity but not complement-dependent cytotoxicity activity. For production of Fab fragments, a truncated heavy-chain (Fd) gene was created by introducing a stop codon near the codon for the amino acid at which papain digestion occurs. Simultaneous expression of the resulting chimeric Fd and light-chain genes in yeast resulted in secretion of properly folded and assembled Fab fragment that bound to target cancer cells.

While a number of single-chain heterologous proteins have been secreted from yeast (1-7), the secretion of foreign multimeric or heterodimeric proteins has not been reported. Of the latter two groups, antibody molecules or the protein fragments that contain their antigen-binding domains, Fab and F(ab')₂ (Fig. 1), are particularly important for a wide variety of applications, including detection and treatment of human disease (8, 9), *in vitro* diagnostic tests (10), and affinity purification methods (11). Chimeric antibodies that consist of mouse variable (V) regions fused to human constant (C) regions may be especially valuable for human therapeutic or *in vivo* diagnostic uses, since they are potentially less immunogenic in humans than are mouse antibodies.

We have described (12) the development of a cDNA cloning strategy for the construction and expression of chimeric antibodies. In this approach, cDNAs coding for the mouse V regions are fused at the immunoglobulin joining (J) region to cDNAs coding for human IgG1 C regions. The cDNA approach for chimeric antibody construction provides an ideal starting point for expression of these genes in microbial systems that either do not undertake mRNA splicing or do so only rarely. In this paper, we describe the secretion from yeast of chimeric antibody and Fab protein. The yeast-secreted molecules (both whole antibody and Fab) bind to target cells as well as their lymphoid cell-derived counterparts. In addition, yeast-derived whole antibody has the same antibody-dependent cellular cytotoxicity (ADCC) activity observed with lymphoid cell-derived whole chimeric antibody but lacks the complement-dependent cytotoxicity (CDC) activity shown by the lymphoid cell-derived antibody.

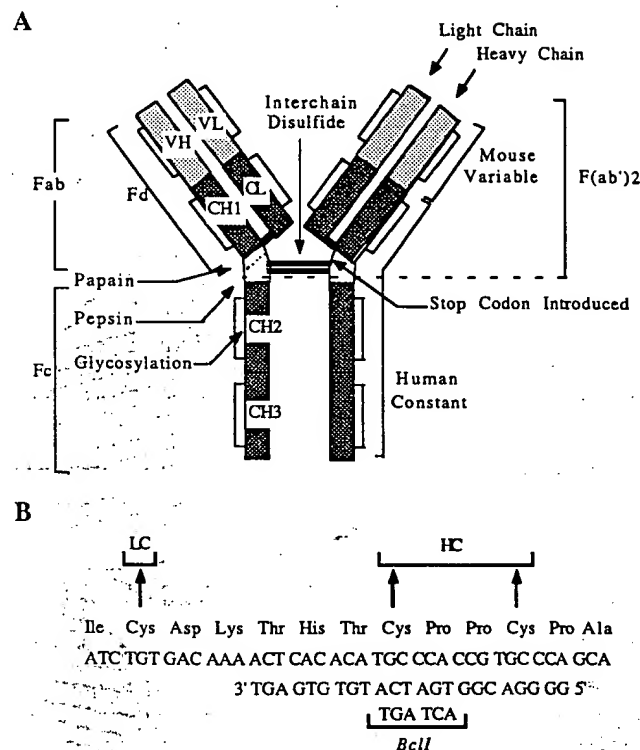


FIG. 1. (A) Structure of mouse-human chimeric IgG1. The locations of papain and pepsin cleavage sites and the structures of F(ab) and F(ab')₂ resulting from digestion with these enzymes are shown. The locations of N-linked glycosylation and a stop codon introduced by site-directed mutagenesis are also shown. V, variable; C, constant; H, heavy; L, light. (B) Site of *in vitro* mutagenesis and DNA sequence of the mutagenesis primer used to place a stop codon and *Bcl* I site in the sequence encoding the hinge region of human γ 1 heavy chain. The *Bcl* I site was converted to *Xho* I by digestion with *Bcl* I followed by treatment with phage T4 DNA polymerase and ligation with *Xho* I linkers. The stop codon was unaffected by this treatment. Arrows indicate interchain disulfide bonds with light chain (LC) and heavy chain (HC).

MATERIALS AND METHODS

Strains and Media. *Escherichia coli* strain MC1061 (13) was used as a host for plasmids. *E. coli* strain 71.18 (14) was used as a host for bacteriophage M13. *Saccharomyces cerevisiae* strain BB331C (*MATa ura3 leu2*) was used as a host for yeast transformations performed as described by Ito *et al.* (15). *E. coli* was grown in TYE broth (1.5% Tryptone/1.0% yeast extract/0.5% NaCl) or agar (1.5% Bacto) supplemented, as

Abbreviations: PGK, phosphoglycerate kinase; ADCC, antibody-dependent cellular cytotoxicity; CDC, complement-dependent cytotoxicity; C, constant; V, variable.

†To whom reprint requests should be addressed.

THIS PAGE BLANK (USPTO)

needed, with ampicillin (50 μ g/ml). Yeast transformants were selected on SD agar (2% glucose/0.67% yeast nitrogen base/2% agar) and grown in SD broth buffered with 50 mM sodium succinate (pH 5.5).

In Vitro Mutagenesis. Site-directed *in vitro* mutagenesis to place restriction sites at yeast or mammalian signal-sequence processing sites and a stop codon in the heavy-chain hinge region was performed as described by Kramer *et al.* (14). Phage plaques containing the desired mutation were identified by plaque-filter hybridization with 32 P-labeled primer.

Enzyme-Linked Immunosorbent Assay (ELISA). Light chain was detected by double-antibody sandwich ELISA (16) using goat anti-human κ antiserum as the coating antibody and peroxidase-labeled goat anti-human κ antiserum for quantitation of bound κ protein. Heavy chain was detected similarly with goat anti-human γ antisera with and without peroxidase label. Association of κ light chains and γ heavy chains was detected with goat anti-human κ antiserum as coating antibody and peroxidase-labeled goat anti-human γ antiserum as the second antibody.

Isolation of Chimeric Whole Antibody and Chimeric Fab from Yeast. Whole antibody was purified from the culture supernatant of a 10-liter fermentation as follows. The culture supernatant was concentrated by ultrafiltration (DC10 ultrafiltration system with spiral cartridge, 30-kDa size cutoff; Amicon), filtered through a 0.45- μ m filter, and concentrated over a YM30 filter (Amicon) to 250 ml. Antibody protein was purified from the concentrated supernatant by protein A-Sepharose chromatography (11). Analysis of this protein by nonreducing polyacrylamide gel electrophoresis followed by Coomassie blue staining and immunoblotting with anti-human κ antiserum (Sigma) as probe revealed a whole immunoglobulin-size protein band against a background smear. The protein in this band was purified by HPLC on an AB_x (5- μ m particle size) column (Baker), with elution by a linear gradient of 10–125 mM potassium phosphate (pH 6.8).

Fab protein was purified from 1 liter of culture supernatant by concentrating over an Amicon YM30 filter, washing with 130 ml of 10 mM potassium phosphate at pH 7.5 (buffer A), and reconcentrating to 12.5 ml. The supernatant was diluted

to 54 ml with buffer A and loaded onto a 1.5-ml S-Sepharose column, washed with 20 ml of buffer A, and eluted with a 40-ml linear gradient of 0–200 mM NaCl in buffer A.

Fab protein prepared by papain digestion (11) of 3-mg of whole L6 mouse antibody or chimeric antibody was purified on a 25-ml S-Sepharose column by elution with an 80-ml linear gradient of 0–120 mM NaCl prepared in 10 mM sodium phosphate (pH 7.5). The Fab protein was eluted at 60 mM NaCl and was free of Fc protein.

Functional Tests of Chimeric Antibody and Fab from Yeast. The following tests were used to assess function: (i) direct binding of whole antibody or Fab to target cells that are positive or negative for the L6 antigen; (ii) competition inhibition of binding of L6 mouse antibody to antigen-positive cells; (iii) ADCC and CDC assays with whole antibody. The binding assays were performed with a Coulter model EPIC-C cell sorter (17). ADCC and CDC assays were performed with 51 Cr-labeled target cells (18, 19) that were exposed to the antibodies and peripheral blood leukocytes or human serum over a period of 4 hr.

RESULTS

Construction of Yeast Expression Plasmids Containing Antibody Genes. To facilitate light- and heavy-chain secretion from yeast, the gene sequences encoding the mature forms of the light and heavy chains of a chimeric anti-tumor antibody (L6, ref. 12) were fused to the yeast invertase signal sequence (20) and placed under the control of the phosphoglycerate kinase (PGK) gene promoter (21). These fusions were then cloned into yeast expression vectors containing the PGK polyadenylation signal (21) to generate pING1441 (light chain, *leu2*) and pING1442 (heavy chain, *ura3*) (Fig. 2A).

The Fd portion of heavy chain consists of the V region and the C_{H1} domain (Fig. 1) and is generated by digestion of whole antibody with papain (11). To enable yeast to synthesize Fd protein, a stop codon was introduced by site-directed mutagenesis of the chimeric heavy-chain gene so that translation would terminate in the hinge region, near the papain recognition site (Fig. 1B). A *Bcl* I site was introduced along

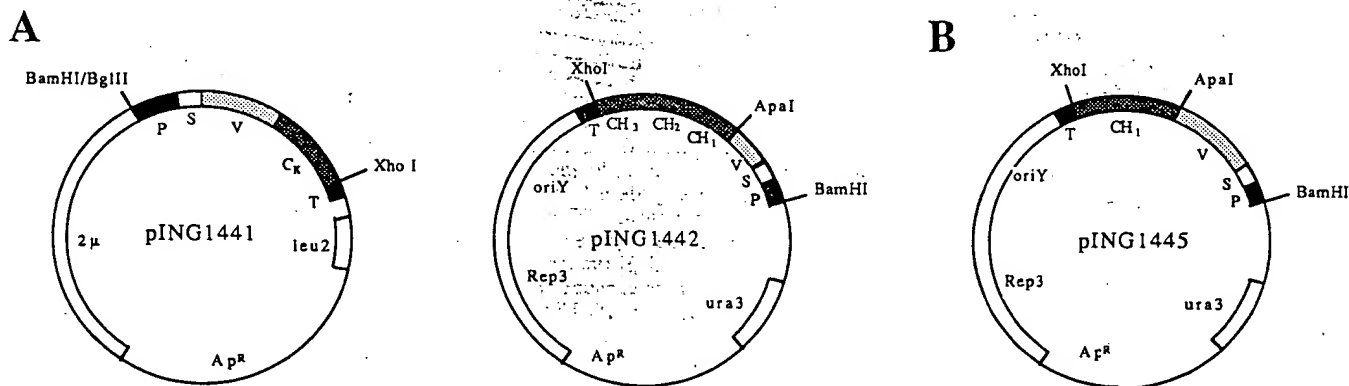


FIG. 2. Structure of yeast immunoglobulin expression plasmids. The fusions of the gene sequences encoding the mature forms of light and heavy chain to the yeast invertase signal sequence and PGK promoter were accomplished by first introducing by *in vitro* mutagenesis a unique restriction site at the signal-sequence processing sites for both the invertase signal sequence (*Pst* I) and the light (*Aat* II) and heavy (*Sst* I) chain genes. These sites were positioned such that a blunt-ended ligation of restriction enzyme-digested, T4 DNA polymerase-treated DNAs resulted in in-phase translational fusions of the 5' end of the mature immunoglobulin chains with the 3' end of the yeast invertase signal sequence. (A) The light-chain expression plasmid, pING1441, was constructed by cloning a *Bgl* II-*Xho* I fragment containing the light-chain gene into a yeast expression vector, pING804CVS (provided by J.-H. Lee, International Genetic Engineering). pING1441 contains the complete 2- μ plasmid (2 μ); the chimeric light-chain gene (V and C_κ regions) fused to the PGK promoter (P), invertase signal sequence (S), and PGK transcription termination and polyadenylation signals (T); and the *leu2* gene as the yeast selective marker. The heavy-chain expression plasmid, pING1442, was constructed by cloning a *Bam* HI-*Xho* I fragment containing the heavy-chain gene into a yeast expression vector, pING1150. pING1442 contains the yeast origin of replication (*oriY*) and a cis-acting stabilization sequence (*Rep3*) from yeast 2- μ plasmid; the chimeric heavy-chain gene (V-region and C-region domains C_{H1}, C_{H2}, and C_{H3}) fused to the PGK promoter, invertase signal sequence, and PGK transcription termination and polyadenylation signals; and the *ura3* gene as the yeast selective marker. (B) The Fd-chain expression plasmid, pING1445, is identical to the heavy-chain expression plasmid, pING1442, with the exception that the *Apal*-*Xho* I restriction fragment in pING1442 encoding C_{H1}, C_{H2}, and C_{H3} has been replaced with a fragment encoding only C_{H1}.

THIS PAGE BLANK (USPTO)

with the stop codon. Following conversion of the *Bcl* I site to *Xho* I, the *Apa* I-*Xho* I C_H1 -region fragment of pING1442 (Fig. 2A) was replaced with an *Apa* I-*Xho* I C_H1 fragment containing the stop codon in the hinge region to generate pING1445 (Fig. 2B).

Secretion of Whole Chimeric Antibody from Yeast. The plasmids pING1441 and pING1442 were cotransformed into *S. cerevisiae* strain BB331C by selection for Ura⁺ Leu⁺ colonies. Ten transformants were grown for 3 days in 5 ml of SD broth and the culture supernatants were analyzed by ELISA for the levels of light chain, heavy chain, and associated light and heavy chains. The culture supernatants of two transformants (nos. 1 and 5) contained light chain at ≈ 100 ng/ml and heavy chain at 50–80 ng/ml, and 50–70% of the heavy chain was associated with light chain. These proteins were concentrated on a Centricon 30 filter (Amicon), electrophoresed in a NaDodSO₄/7% polyacrylamide gel under nonreducing conditions, and transferred to nitrocellulose. κ -crossreactive protein was detected with goat anti-human κ antiserum followed by peroxidase-labeled rabbit anti-goat antiserum. A faint, but distinct, band that comigrated with purified chimeric L6 antibody produced by Sp2/0 cells was observed in the lanes containing both supernatants (data not shown). These results suggested that the yeast transformants were synthesizing and secreting a fully assembled chimeric antibody that was very similar to lymphoid cell-derived antibody.

To prepare sufficient quantities of yeast-derived antibody for detailed characterization of structure and function, transformant no. 5 was grown in a 10-liter fermentor for 58 hr. Whole antibody was purified from this culture medium as described in *Materials and Methods*. ELISA analysis of column fractions from HPLC (AB_x 5- μ m) revealed a heavy (γ) and light (κ) chain-crossreactive peak corresponding to a distinct A₂₈₀ peak. Analysis of these fractions by nonreducing NaDodSO₄/polyacrylamide gel electrophoresis revealed a protein that comigrated with L6 chimeric antibody purified from Sp2/0 cells (Fig. 3). Under reducing conditions, the protein in these fractions was resolved into a light-chain band, which comigrated with the light chain of L6 chimeric antibody from Sp2/0 cells, and two heavy chain bands, which migrated near the heavy chain from Sp2/0 cells (Fig. 3). This purified preparation of yeast-produced chimeric antibody was used in further binding and function assays.

Secretion of Chimeric Fab from Yeast. The plasmids pING1441 (Fig. 2A) and pING1445 (Fig. 2B) were cotransformed into *S. cerevisiae* BB331C and the transformants were grown in broth under selective conditions as described above. The culture supernatants were assayed by ELISA and were found to contain light chain at 100–200 ng/ml. To determine whether the cells secreted a Fab-size protein, the culture supernatants were concentrated with Centricon 30 filters, electrophoresed in a nonreducing NaDodSO₄/10% polyacrylamide gel, and electrophoretically transferred to nitrocellulose paper. Light chain-crossreactive protein was detected with goat anti-human κ antiserum followed by peroxidase-labeled rabbit anti-goat antiserum. One of the five transformants secreted a distinct κ -crossreactive band, which migrated at the expected position for Fab protein (data not shown).

The yeast strain that secreted the Fab-size protein was grown in 1 liter of SD broth for 4 days at 30°C and Fab protein was purified from the culture supernatant. Nonreducing NaDodSO₄/polyacrylamide gel analysis of pooled S-Sepharose column fractions containing anti-human κ -crossreactive protein revealed a 46-kDa protein, comigrating with chimeric Fab prepared by papain digestion of Sp2/0 cell-derived chimeric whole antibody (Fig. 4). Electrophoresis under reducing conditions resolved this protein into two bands that migrated at 23 and 25 kDa (Fig. 4). The 23-kDa

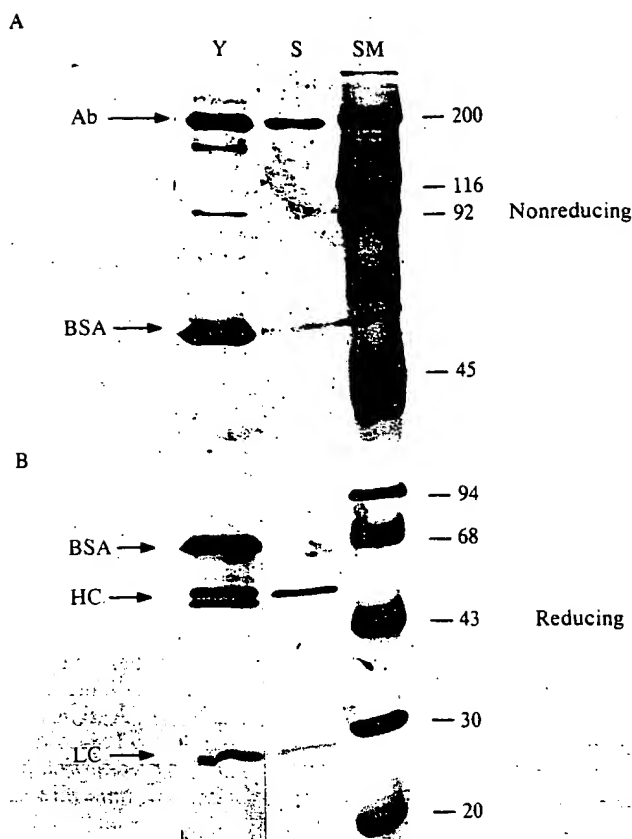


FIG. 3. Silver-stained NaDodSO₄/polyacrylamide gels showing purified chimeric antibody secreted by yeast. (A) Nonreducing 7% gel. (B) Reducing 10% gel. Intensely stained band at 68 kDa on both gels is bovine serum albumin (BSA), which was present as a carrier protein. Size marker (lanes SM) molecular masses (in kDa at right) and relevant yeast-derived (lanes Y) and Sp2/0-derived (lanes S) antibody bands [Ab on nonreducing gel; HC (heavy chain) and LC (light chain) on reducing gel] are identified.

band was identified as light chain by its reactivity on an immunoblot with anti-human κ antiserum. These results were consistent with the predicted molecular masses, based on nucleotide sequence, for fully processed L6 chimeric light chain (23.3 kDa) and Fd chain (24.8 kDa).

Binding Characteristics of Chimeric Whole Antibody and Fab Protein Secreted by Yeast. The purification from yeast culture supernatants of protein of the expected size of whole antibody and Fab suggested that the yeast transformants were secreting correctly folded, functional molecules. This hypothesis was confirmed by performing direct and competition binding assays with a human colon carcinoma cell line (line C-3347) that expresses 5×10^5 molecules of the L6 tumor antigen per cell (17). In the direct binding assay, both whole antibody and Fab from yeast bound to the target cancer cells but not to a control cell line that lacked the L6 antigen (data not shown). In the competition assay using mouse L6 antibody, 50% binding inhibition was observed at the same concentration (2 μ g/ml) for both yeast-derived and Sp2/0 cell-derived whole chimeric antibody (Fig. 5). In the same assay, the yeast-derived L6 chimeric Fab behaved identically to both Sp2/0 cell-derived chimeric and mouse Fab proteins prepared by papain digestion (Fig. 5). Fifty percent inhibition of mouse L6 antibody was achieved by the yeast-derived Fab at 7 μ g/ml.

Chimeric Whole Antibody from Yeast Mediates ADCC but Not CDC. Two additional tests for function were performed with chimeric whole antibody from yeast: (i) mediation of ADCC in the presence of human peripheral blood leukocytes

THIS PAGE BLANK (USPTO)

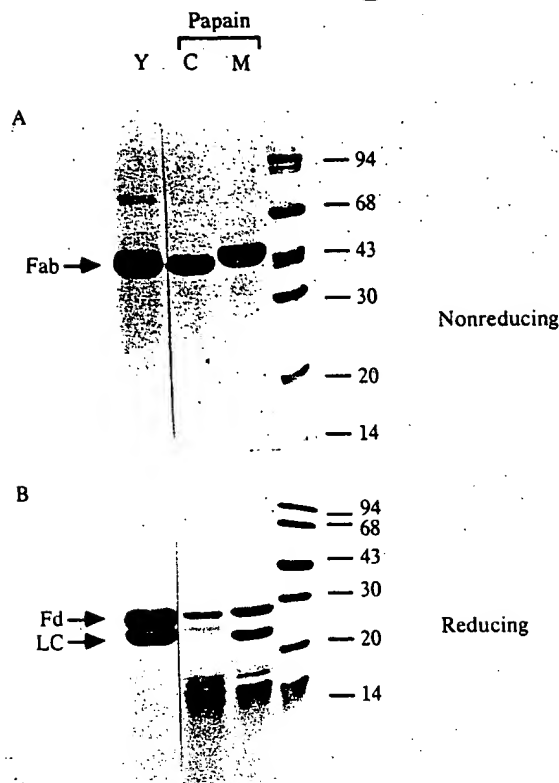


FIG. 4. Coomassie blue-stained gel comparing purified chimeric Fab protein secreted by yeast (lanes Y) with chimeric (lanes C) and mouse (lanes M) Fab fragments produced by papain digestion of intact antibody. (A) Nonreducing 10% gel. (B) Reducing 12% gel. Size marker molecular masses (kDa) and relevant bands are shown. LC, light chain.

and (ii) mediation of tumor-cell lysis in the presence of human complement. ADCC activity of yeast-derived chimeric L6 antibody was slightly higher than that of Sp2/0 cell-derived chimeric antibody, and the ADCC activities of both chimeric antibodies were higher than that of the mouse L6 antibody (Table 1). Yeast-derived L6 chimeric antibody failed to mediate CDC, even at the highest antibody concentration, whereas Sp2/0 cell-de-

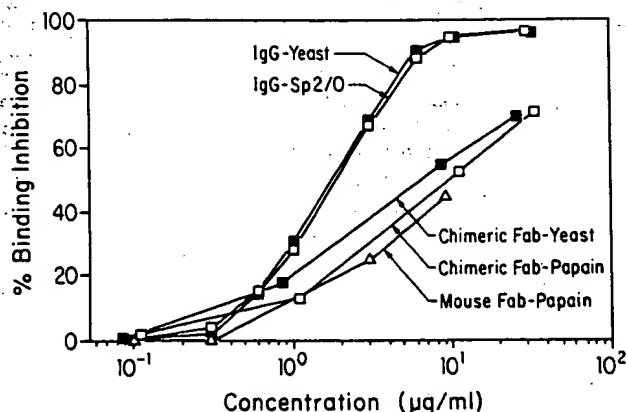


FIG. 5. Comparison in antibody-competition binding assays between whole chimeric L6 antibody (IgG) derived from yeast and Sp2/0 cells, chimeric L6 Fab derived from yeast or prepared by papain digestion of whole chimeric antibody isolated from Sp2/0 cells and mouse Fab prepared by papain digestion of L6 antibody. C-3347 colon carcinoma cells were incubated with various concentrations of unlabeled blocking antibodies before addition of fluorescein isothiocyanate-conjugated mouse L6 antibody (3 µg/ml). Inhibition was measured by flow cytometry.

Table 1. ADCC analysis

Antibody	Conc., µg/ml	% cytotoxicity
Standard mouse L6	5.0	42
	1.0	48
	0.1	96
Sp2/0 chimeric L6	1.0	96
	0.1	71
	0.01	54
	0.001	37
Yeast chimeric L6	1.0	114
	0.1	108
	0.01	76
	0.001	60
None	0	23

The colon carcinoma target cells (line C-3347) were labeled with ⁵¹Cr and exposed for 4 hr to a combination of monoclonal antibody and human peripheral blood leukocytes (100 per target cell), and the release of ⁵¹Cr was measured subsequently. The release of ⁵¹Cr (after corrections of values for spontaneous release from untreated cells) is a measure of cytotoxicity.

Yeast-derived chimeric L6 antibody and mouse L6 antibody exhibited the expected cytotoxic activity (Table 2).

DISCUSSION

We have engineered the yeast *S. cerevisiae* to secrete functional mouse-human chimeric antibody and Fab protein into the culture medium. This was accomplished by simultaneous expression of the mature light-chain gene and the heavy-chain gene or a truncated heavy-chain (Fd) gene fused to the yeast invertase signal sequence and the PGK promoter and polyadenylation signal. Several lines of evidence support the thesis that these proteins are correctly folded. (i) Proteins of the expected size for whole antibody and Fab were purified from the culture supernatants of yeast cells expressing the chimeric light- and heavy- or Fd-chain genes (Figs. 3 and 4). (ii) The whole antibody and Fab from yeast behaved indistinguishably from their lymphoid cell-derived counterparts in both direct and competition binding assays (Fig. 5). (iii) The chimeric whole antibody from yeast exhibited the same ADCC activity as the chimeric antibody from Sp2/0 cells (Table 1).

There have been a number of reports of secretion from yeast of heterologous proteins fused to yeast signal sequences (2, 4, 5, 7). All of the proteins in these examples were composed of a single polypeptide chain. Although there is one example of functional mouse antibody (IgM) production in yeast, the antibody in this case was found only intracellularly in vacuoles; only unassociated light and heavy chains were detected in the culture supernatant (22).

Table 2. CDC analysis

Antibody	Conc., µg/ml	Complement*	% cytotoxicity†
Standard mouse L6	5	+	122
	1	+	53
	5	-	1
Sp2/0 chimeric L6	5	+	73
	1	+	22
	0.1	+	5
	5	-	2
Yeast chimeric L6	5	+	3
	1	+	2
	0.1	+	4

*Human serum from a healthy subject was used as the source of complement.

†⁵¹Cr-labeled C-3347 cells were exposed to human complement and antibody. CDC was measured by a 4-hr ⁵¹Cr-release assay.

THIS PAGE BLANK (USPTO)

The chimeric whole antibody from yeast mediated ADCC in an identical fashion to Sp2/0 cell-derived chimeric antibody (Table 1). These results suggest that the yeast-derived antibody is equivalent to the lymphoid cell-derived antibody in interacting with Fc receptors on killer cells and activating them to mediate ADCC. Interestingly, the chimeric mouse-human antibody from either lymphoid cells or yeast was more efficient at ADCC than was the mouse antibody. Thus, the Fc-receptor interaction involved in ADCC appears to be primarily determined by the amino acid sequence of the Fc portion of the antibody and is probably not affected by the altered glycosylation patterns expected for yeast-derived antibody. This conclusion is consistent with the observation that binding of Fc receptor type I by IgG, which may play a central role in ADCC (23), occurs in the region linking the C_{H2} domain to the hinge (24). This region of the protein is an exposed, flexible strand (24) and is well removed from the site of N-linked glycosylation (25).

An intriguing result is that the yeast-derived antibody lacks the ability to activate complement to lyse target cells (Table 2). This may reflect differences in the glycosylation patterns of the yeast-derived and Sp2/0 cell-derived chimeric antibodies. Several observations support this hypothesis. Binding of complement component C1q, which initiates CDC, occurs within the C_{H2} domain of human IgG (26), which is also the region of N-linked glycosylation. Second, elimination of the N-linked glycosylation site in mouse IgG2b by *in vitro* mutagenesis of asparagine-297 to alanine results in reduced affinity of the antibody for human C1q and complete loss of CDC (27). Third, while yeast and mammalian cells recognize the same peptide signal [Asn-Xaa-(Ser/Thr)] for glycosylation and utilize a similar pathway for core oligosaccharide synthesis in the endoplasmic reticulum (28), the type and extent of outer chain glycosylation appear to be quite different in these cell types (29). Indeed, comparison of yeast- and Sp2/0 cell-derived heavy chains by reducing NaDod-SO₄/polyacrylamide gel electrophoresis revealed differences in protein mobility (Fig. 4), which may be caused by differences in glycosylation. Such glycosylation differences may be sufficient to cause a loss of CDC activity. Further studies will be required to establish the exact cause of these size differences and their possible relationship to CDC activity.

Fab proteins may be especially useful for certain diagnostic and therapeutic applications, including *in vivo* tumor imaging (8) and drug or toxin delivery to tumors. Mouse-human chimeric Fab proteins are particularly attractive for these uses because they may be less immunogenic than mouse Fab proteins in humans. Current methods for Fab production involve papain digestion of purified whole antibody (11). This approach can be problematic, since not all antibodies are equivalent in their susceptibility to papain cleavage (11) and additional purification steps are required beyond that for the whole antibody. In addition, papain digestion can result in partial degradation of the Fab (Fig. 4). By contrast, direct production of Fab by yeast or, as described elsewhere, by *E. coli* or Sp2/0 cells (30) can yield a highly purified, homogeneous Fab preparation that has the same binding activity as the Fab prepared by papain digestion (Fig. 5). Further, the introduction by site-directed mutagenesis of a restriction site in conjunction with the stop codon in the Fd sequence permits manipulations at the 3' end of the gene that can yield Fab proteins with altered properties, such as enhanced affinity in labeling reactions, or allow direct production of Fab molecules linked to various proteins.

We thank Bob Haselbeck and Pam Smith for excellent technical assistance. We also thank Ingegerd Hellstrom and coworkers for

performing binding and activity assays, Yun-Long Lin for performing yeast fermentations, and Lindley C. Blair, Arup Sen, Randy Wall, and Gary Wilcox for useful comments on the manuscript. The work was supported by International Genetic Engineering, Inc.

- Hitzeman, R. A., Leung, D. W., Perry, L. J., Kohr, W. J., Levine, H. L. & Goeddel, D. V. (1983) *Science* 219, 620-625.
- Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlein, U. A., Masiarz, F. R., Mullenbach, G. T., Urden, M. S., Valenzuela, P. & Barr, P. J. (1984) *Proc. Natl. Acad. Sci. USA* 81, 4642-4646.
- Innis, M. A., Holland, M. J., McCabe, P. C., Cole, G. E., Wittman, V. P., Tal, R., Watt, W. K., Gelfand, D. H., Holland, J. P. & Meade, J. H. (1985) *Science* 228, 21-26.
- Smith, R. A., Duncan, M. J. & Moir, D. T. (1985) *Science* 229, 1219-1224.
- Chang, C. N., Matteucci, M. L., Perry, J., Wulf, J. J., Chen, C. Y. & Hitzeman, R. A. (1986) *Mol. Cell. Biol.* 6, 1812-1819.
- Jigami, Y., Muraki, M., Harada, N. & Tanada, H. (1986) *Gene* 43, 273-279.
- Moir, D. T. & Dumais, D. R. (1987) *Gene* 56, 209-217.
- Delaloy, B., Bischof-Delaloye, A., Buchegger, F., Von Fliedner, V., Grob, J.-P., Volant, J.-C., Pettavel, J. & Mach, J.-P. (1986) *J. Clin. Invest.* 77, 301-311.
- Levy, R. & Millre, R. A. (1983) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* 42, 2650-2656.
- Tam, M. R., Goldstein, L. C. & Yelton, D. E. (1985) in *Manual of Clinical Microbiology*, eds. Lennette, E. H., Balows, A., Hausler, W. J., Jr., & Shadomy, H. J. (Am. Soc. Microbiol. Washington, DC), 4th Ed., pp. 905-909.
- Johnstone, A. & Thorpe, R. (1982) *Immuno Chemistry in Practice* (Blackwell, Oxford), pp. 52-53.
- Liu, A. Y., Robinson, R. R., Hellstrom, K. E., Murray, E. D., Jr., Chang, C. P. & Hellstrom, I. (1987) *Proc. Natl. Acad. Sci. USA* 84, 3439-3443.
- Casadaban, M. J. & Cohen, S. N. (1980) *J. Mol. Biol.* 138, 179-207.
- Kramer, W., Drutsa, V., Jansen, H.-W., Kramer, B., Pflugfelder, M. & Fritz, H.-J. (1984) *Nucleic Acids Res.* 12, 9441-9456.
- Ito, H., Fukuda, Y., Murata, K. & Kimura, A. (1983) *J. Bacteriol.* 153, 163-168.
- Voller, A., Bidwell, D. & Bartlett, A. (1980) in *Manual of Clinical Immunology*, eds. Rose, N. R. & Friedman, H. (Am. Soc. Microbiol., Washington, DC), 2nd Ed., pp. 359-371.
- Hellstrom, I., Horn, D., Linsley, P., Brown, J. P., Brankovan, V. & Hellstrom, K. E. (1986) *Cancer Res.* 46, 3917-3923.
- Hellstrom, I., Brankovan, V. & Hellstrom, K. E. (1985) *Proc. Natl. Acad. Sci. USA* 82, 1499-1502.
- Hellstrom, I., Beaumier, P. L. & Hellstrom, K. E. (1986) *Proc. Natl. Acad. Sci. USA* 83, 7059-7063.
- Taussig, R. & Carlson, M. (1983) *Nucleic Acids Res.* 11, 1943-1954.
- Hitzeman, R. A., Hagie, F. E., Hayflick, J. S., Chen, C. Y., Seeburg, P. H. & Derynck, R. (1982) *Nucleic Acids Res.* 10, 7791-7808.
- Wood, C. R., Boss, M. A., Kenten, J. H., Calvert, J. E., Roberts, N. A. & Emtage, J. S. (1985) *Nature (London)* 314, 446-449.
- Unkeless, J. C., Scigliano, E. & Freedman, V. H. (1988) *Annu. Rev. Immunol.* 6, 251-281.
- Duncan, A. R., Woof, J. M., Partridge, L. J., Burton, D. R. & Winter, G. (1988) *Nature (London)* 332, 563-564.
- Jeske, D. J. & Capra, J. D. (1984) in *Fundamental Immunology*, ed. Paul, W. E. (Raven, New York), pp. 131-165.
- Yasmeen, D., Ellerson, J. R., Dorrington, K. J. & Painter, R. (1976) *J. Immunol.* 116, 1664-1667.
- Duncan, A. R. & Winter, G. (1988) *Nature (London)* 332, 738-740.
- Schekman, R. (1985) *Annu. Rev. Cell Biol.* 1, 115-143.
- Kukuruzinska, M. A., Bergh, M. L. E. & Jackson, B. J. (1987) *Annu. Rev. Biochem.* 56, 915-944.
- Better, M., Chang, C. P., Robinson, R. R. & Horwitz, A. H. (1988) *Science* 240, 1041-1043.

THIS PAGE BLANK (USPTO)

FALL 1988 B10L. LIB.

STANFORD UNIVERSITY
LIBRARIES

NOV 30 1988

TAL
NOVEMBER 1988

VOLUME 85

NUMBER 22



Proceedings OF THE National Academy of Sciences

OF THE UNITED STATES OF AMERICA

THIS PAGE BLANK (USPTO)

14. P. T. Jones, P. H. Dear, J. Foote, M. S. Neuber, G. Winter, *ibid.* 321, 522 (1986).
15. M. Potter, *Adv. Immunol.* 25, 141 (1977); R. M. Perlmutter *et al.*, *ibid.* 35, 1 (1984); S. Rudikoff and M. Potter, *Biochemistry* 13, 4033 (1974); S. Rudikoff, Y. Satow, E. Padlan, D. Davies, M. Potter, *Mol. Immunol.* 18, 705 (1981).
16. D. M. Segal *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 71, 4298 (1974).
17. L. Hendershot, D. Bole, J. F. Kearney, *Immunol. Today* 8, 111 (1987).
18. M. S. Briggs and L. M. Gierasch, *Adv. Protein Chem.* 38, 109 (1986).
19. N. R. Mowla, K. Nakamura, M. Inouye, *J. Biol. Chem.* 255, 27 (1980).
20. H. Inouye, W. Barnes, J. Beckwith, *J. Bacteriol.* 149, 434 (1982).
21. The two chains of the affinity-purified F_v protein were separated using SDS-PAGE and blotted onto an activated glass fiber sheet (22). The two bands corresponding to each of the two chains were each individually subjected to six cycles of gas-phase sequencing. In both sequenator profiles, the NH₂-terminal amino acids expected from precise cleavage of both signals were seen exclusively.
22. C. Eckerskorn, W. Mewes, H. Goretzki, F. Lottspeich, *Eur. J. Biochem.*, in press.
23. H. Metzger, B. Chesebro, N. M. Hadler, J. Lee, N. Otchin, in *Progress in Immunology: Proceedings of the 1st Congress of Immunology* (Academic Press, New York, 1971), pp. 253-267.
24. G. Schumacher, D. Sizmann, H. Haug, P. Buckel, A. Böck, *Nucleic Acids Res.* 14, 5713 (1986).
25. J. R. Hobbs, *Essays Med. Biochem.* 1, 105 (1975).
26. D. Inbar, J. Hochman, D. Givol, *Proc. Natl. Acad. Sci. U.S.A.* 69, 2659 (1972).
27. J. Sen and S. Beychok, *Proteins* 1, 256 (1986).
28. R. Huber and W. S. Bennett, *Nature* 326, 334 (1987); M. Marquart and J. Deisenhofer, *Immunol. Today* 3, 160 (1982).
29. C. Yanisch-Perron *et al.*, *Gene* 33, 103 (1985).
30. A. Skerra and A. Plückhuhn, in preparation.
31. S. P. Fling and D. S. Gregerson, *Anal. Biochem.* 155, 83 (1986).

We monitored the selectivity of the cell fractionation using β -lactamase as a periplasmic and β -galactosidase as a cytoplasmic marker enzyme. Under the conditions described, approximately 90% of total β -lactamase activity and less than 0.5% of total β -galactosidase activity were found in the periplasmic fraction.

33. B. Chesebro and H. Metzger, *Biochemistry* 11, 766 (1972).
34. R. K. Scopes, *Protein Purification: Principles and Practice* (Springer, New York, 1982), pp. 241-242.
35. G. Scatchard, *Ann. N.Y. Acad. Sci.* 51, 660 (1949).
36. We thank C. Eckerskorn and F. Lottspeich for the NH₂-terminal protein sequencing and R. Glockshuber for the preparation of the phosphorylcholine affinity column. Supported by grant BCT0372 from the Bundesministerium für Forschung und Technologie to A.P., and by a predoctoral fellowship from the Stiftung Volkswagenwerk and the Fonds der Chemischen Industrie to A.S.

7 January 1988; accepted 18 March 1988

Escherichia coli Secretion of an Active Chimeric Antibody Fragment

MARC BETTER, C. PAUL CHANG, RANDY R. ROBINSON, ARNOLD H. HORWITZ

A chimeric mouse-human Fab protein that binds specifically to the human carcinoma cell line C3347 has been expressed and secreted from *Escherichia coli*. This molecule, which contains functionally assembled kappa and Fd proteins, binds as effectively to sites on the surface of C3347 cells as Fab fragments prepared proteolytically from whole chimeric or mouse antibody. The production in *Escherichia coli* of foreign heterodimeric protein reagents, such as Fab, should prove useful in the management of human disease.

THE BINDING-SELECTIVITY OF ANTIBODY molecules makes them suited for applications as diverse as affinity chromatography, diagnostic reagents, and therapeutics in the detection and treatment of human diseases. Monoclonal antibodies are especially useful for these purposes because they can be prepared with homogeneous recognition specificities targeted at virtually an unlimited number of antigenic determinants. The protein domains that confer these antigen recognition determinants can be proteolytically separated from the remainder of the molecule and still retain their antigen-binding ability. This portion of an antibody (Fab) is roughly one-third the size of an intact immunoglobulin G (IgG) (about 48 kD) and exhibits monovalent antigen binding. The similar F(ab')₂ portion retains divalent antigen-binding capacity and contains both recognition domains linked by two interchain disulfide bridges. Antibodies differ, however, in their susceptibility to proteolytic cleavage, and preparations can be heterogeneous. The relatively simple structure of a Fab (5 disulfide

bonds) compared to an intact antibody (16 disulfide bonds) and the therapeutic usefulness of Fab molecules make them attractive targets for production by microbial fermentation after appropriate protein engineering. Here we discuss the expression of a mouse-human chimeric Fab in *Escherichia coli*, that is, a molecule that contains the variable regions (antigen recognition domains) from a mouse monoclonal antibody and the C_H1 and C_κ constant regions from a human IgG1 antibody.

Each protein chain of a Fab has two intrachain disulfide bonds that stabilize functional domains, and a single cysteine involved in interchain disulfide linkage. *Escherichia coli* has been used to produce individual immunoglobulin chains internally that are not properly folded (1, 2), or individually secreted chains (3); however, for *E. coli* to assemble the truncated heavy chain (Fd) and κ into the correct heterodimeric molecule, both chains must be translated simultaneously and secreted. This operation would then mimic the cognate immunoglobulin assembly process.

The chimeric L6 antibody (4), directed toward a ganglioside antigen expressed on the cell surface of many human carcinomas

(5), has been described. This antibody, prepared from the culture supernatant of an Sp2/0 transfectoma cell line, is expressed from the cDNA copies of the chimeric L6 IgG γ and κ chain genes. The cDNA clones for these two chimeric genes were used as the starting point for expression of a Fab molecule in mammalian cells and bacteria.

A termination codon was introduced into the chimeric heavy chain gene at amino acid 228 by site-directed mutagenesis (6) (Fig. 1a) in a manner that introduced a Bcl I restriction site. A similar step introduced a restriction site, Sst I, into the coding region at the processing cleavage site of the native heavy chain leader peptide and the mature heavy chain. Site-directed mutagenesis was

Table 1. Binding activity of bacterial Fab to human carcinoma cells. Target cells were incubated for 30 min at 4°C with each antibody or Fab, followed by incubation with fluorescein isothiocyanate (FITC)-labeled goat antibody against human κ for the bacterial Fab, FITC-labeled goat antibody against mouse IgG for the L6 mouse antibody, FITC-labeled goat antibody against mouse κ for L6 mouse Fab, or goat antibody against human IgG for the chimeric L6 antibody. We determined antibody binding to the cell surface by using a Coulter model EPIC-C cell sorter. FITC-labeled antibodies were obtained from TAGO.

Antibody	Binding ratio*	
	C3347 cells L6+	T51 cells L6-
Mouse L6	95	1
Sp2/0 chimeric L6	116	1
Bacterial L6 Fab	54	1
Mouse L6 Fab†	16	1

*The binding ratio is the number of times brighter a test sample is than a control sample treated with FITC-conjugated second antibody. Quantitative differences in binding to C3347 cells probably reflect the relative activity of individual FITC-conjugates. Data shown are from one of two similar binding assays. †Prepared by enzymatic digestion of mouse L6 antibody with papain.

International Genetic Engineering, Inc. (INGENE), 1545 17th Street, Santa Monica, CA 90404.

THIS PAGE BLANK (USPTO)

also used to insert an Aat II restriction site into the chimeric κ chain gene at the junction of the leader peptide and the mature processed κ chain. Each coding sequence, Fd and κ , was fused to the leader peptide segment of the bacterial *pelB* gene (pectate lyase) from *Erwinia carotovora* (7) generating a gene fusion between the leader peptide segment of the pectate lyase gene and the mature coding sequence of the Fd and κ genes. This bacterial leader sequence was chosen to direct membrane translocation in *E. coli* since pectate lyase can be secreted to high levels under the control of a regulated promoter. To ensure that both Fd and κ were also translated in close physical proximity, we assembled a plasmid that codes both genes in a single dicistronic message (Fig. 1b). This operon was placed under the control of the inducible *araB* promoter from *Salmonella typhimurium* (8) and expressed in *E. coli*.

Examination of culture supernatants or extracts of the periplasmic space (9) of *E. coli* by enzyme-linked immunosorbent assay (ELISA) for chimeric κ with antibody against human κ (Cappel), or Fab production with antibody against human Fd (Calbiochem) and antibody against human κ , revealed that about 90% of the secreted κ chain accumulated in the culture medium. This was a surprising observation that allowed simple purification of this material from induced bacterial cultures. Approximately 2 mg/liter of material reactive as Fab in an ELISA is secreted into culture super-

natants of MC1061 (pIT106). Immunoblot analysis with antibody against human κ revealed that under nonreducing gel conditions, the predominant reactive species had a molecular size of about 48 kD. Under gel conditions where disulfide linkages were reduced, the predominant species had a molecular size of about 23 kD. These observations are consistent with the predicted molecular sizes of the processed chimeric Fab (κ , 23.3 kD; Fd, 24.7 kD), and suggest that the material is properly assembled.

For purification of Fab, bacterial supernatants were concentrated, filtered, and loaded on an SP disk equilibrated with 10 mM phosphate buffer, pH 7.5. Fab was eluted with 0.2M NaCl and purified by S Sepharose column chromatography, where it was eluted as a single peak with a linear 0 to 0.12M NaCl gradient. The immunologically reactive material was more than 90% pure as determined by SDS-polyacrylamide gel electrophoresis (PAGE) followed by Coomassie blue staining (Fig. 2). Purified material has a molecular size of about 48 kD under nonreducing gel conditions and about 24.5 and 23 kD under reducing gel conditions. The 23-kD band is immunologically distinguishable by using antibody against human κ .

Purified bacterial Fab was tested for binding to L6 antigen-containing cells (Table 1). Bacterial Fab bound specifically to the human colon carcinoma cell line C3347; cells from the T cell line T51 served as a negative control. Bacterially produced Fab

also exhibited characteristic binding inhibition of FITC-labeled mouse L6 antibody to the surface of antigen-positive C3347 colon carcinoma cells (Fig. 3). We tested bacterially produced chimeric Fab, proteolytically prepared Fab from L6 mouse antibody, L6 chimeric antibody, and Fab prepared from Sp2/0 cells transfected with the truncated chimeric Fd and the chimeric κ gene. All Fab preparations have essentially identical binding inhibition profiles. The proteolytically produced Fab contains a significant proportion of degraded, low molecular size peptides, whereas chimeric Fab from bacteria or Sp2/0 cells is homogeneous (Fig. 2).

Protein engineering allowed the expression in *E. coli* of a functional chimeric Fab that has binding specificity for a human carcinoma cell marker. The finding that *E. coli* can be engineered to secrete a foreign heterodimeric molecule builds on the earlier report that intrachain disulfide bonds can form correctly in proteins secreted into the periplasmic space of *E. coli* (10). One useful application for bacterially produced Fab molecules will be in tumor imaging in vivo (11, 12) [tumor marker-specific Fabs produced proteolytically from whole antibodies have already been used for this purpose (13-15)]. A great advantage of the engineered bacterial Fab is that the protein heterogeneity that results from nonspecific cleavages and differences in the susceptibility of anti-

Fig. 1. (a) Strategy for introduction of a stop codon and Bcl I restriction site at amino acid 228 in the chimeric L6 heavy chain cDNA clone. An M13mp19 subclone containing a portion (V_H , C_H1 , C_H2) of the chimeric L6 γ gene, pING1400, was used as a substrate for site-directed mutagenesis with the primer shown. Arrows indicate the cysteine residue involved in disulfide linkage with light chain (LC) and those involved in linkage with heavy chain (HC) in an intact IgG. **(b)** Map of plasmid pIT106 and a schematic view of the relevant Fab operon placed under the control of the *S. typhimurium araB* promoter (*ParaB*). The 300-bp *pelB* leader sequence cassette including the 22-amino acid *pelB* leader and 230 bp of upstream DNA was derived from pSS1004 (7). Gene fusions were constructed with *pelB* and both Fd and κ genes. *Erwinia carotovora* DNA sequences 5' to the -48 position were removed by means of an Nde I restriction site, and the coding sequences for these genes were assembled [in the *araB* expression plasmid pIT2 (17)] into a single transcription unit with 70 bp separating the two genes. The pectate lyase ribosome binding site is located adjacent to each gene. In addition, the DNA sequence from -48 to the *pelB*:antibody gene junction is a direct repeat preceding the Fd and κ genes. The 28-bp *trpA* transcriptional terminator (Pharmacia) was cloned at the end of the transcription unit (Tm). Restriction endonuclease sites are as indicated: P, Pst I; Bg, Bgl II; Bc, Bcl I; and H, Hind III. The solid area denotes the approximate location of the *pelB* leader sequence.

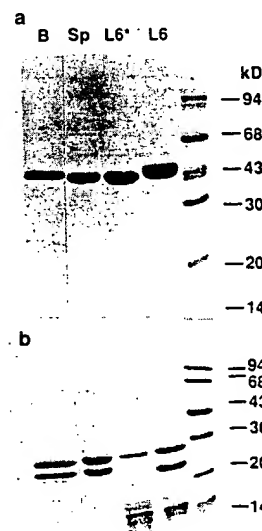
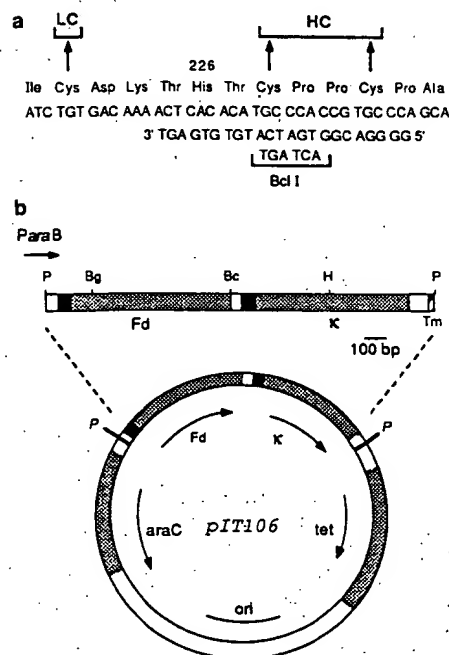
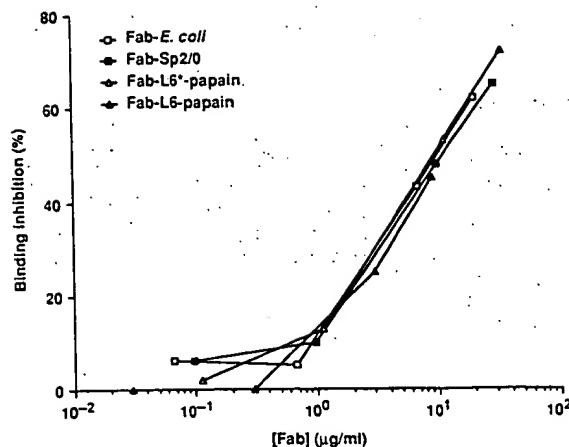


Fig. 2. SDS-PAGE comparison of bacterially (B) and Sp2/0-produced (Sp) Fab to papain-produced mouse L6 and chimeric L6* Fab. Mouse L6 and chimeric L6 antibodies were digested with papain (18), and Fab was purified by S-Sepharose chromatography. Each protein was examined by SDS-PAGE on a 10% gel under nonreducing conditions (a) and on a 12% gel under reducing conditions (b).

THIS PAGE BLANK (USPTO)

Fig. 3. Binding inhibition of bacterial Fab. Bacterial Fab, Sp2/0 Fab, and proteolytically produced chimeric L6 Fab (L6*) and mouse L6 Fab were used to inhibit FITC-labeled mouse L6 antibody binding to the surface of antigen-positive C3347 colon carcinoma cells.



bodies to protease cleavage will be obviated; a consistent, homogeneous preparation can be produced. Of additional interest is the relative ease with which the Fab cDNA genes can be modified before expression in bacteria. For example, modifications of the primary structure of either the Fd or κ chain (or both) that are useful for subsequent conjugation of imaging or therapeutic agents or fusion to other peptides (16) can be introduced by site-directed mutagenesis

techniques. We found that *E. coli* can properly assemble a functional two-protein unit with a complex pattern of intra- and inter-chain disulfide linkages and that sufficient quantities of this material may be prepared for eventual use as a human diagnostic and therapeutic reagent.

REFERENCES AND NOTES

1. M. A. Boss, J. H. Kenten, C. R. Wood, J. S. Emtage, *Nucleic Acids Res.* 12, 3791 (1984).
2. S. Cabilly *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 81,

- 3273 (1984).
3. O. Zemel-Dreazen and A. Zamir, *Gene* 27, 315 (1984).
4. A. Y. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 84, 3439 (1987).
5. I. Hellstrom *et al.*, *Cancer Res.* 46, 3917 (1984).
6. W. Kramer *et al.*, *Nucleic Acids Res.* 12, 9441 (1984).
7. S. P. Lei, H. C. Lin, S. S. Wang, J. Callaway, G. Wilcox, *J. Bacteriol.* 169, 4379 (1987).
8. S. Johnston *et al.*, *Gene* 34, 137 (1985).
9. N. Yanagida, T. Uozumi, T. Beppu, *J. Bacteriol.* 166, 937 (1986).
10. H. M. Hsiung, N. G. Mayne, G. W. Becker, *Biotechnology* 4, 991 (1986).
11. E. Haber, *Annu. Rev. Med.* 37, 249 (1986).
12. J. R. Zalcberg, *Am. J. Clin. Oncol.* 8, 481 (1985).
13. B. Delaloye *et al.*, *J. Clin. Invest.* 77, 301 (1986).
14. J. F. Eary *et al.*, in preparation; P. G. Abrams *et al.*, *Second International Conference of Monoclonal Antibody Immunoconjugates for Cancer* (University of California, San Diego Cancer Center, San Diego, 1987), p. 12.
15. R. L. Wahl, D. W. Parker, G. W. Philpott, *J. Nucl. Med.* 24, 316 (1983).
16. E. S. Vitetta *et al.*, *Science* 238, 1098 (1987).
17. L. Masson and D. S. Ray, *Nucleic Acids Res.* 14, 5693 (1986).
18. A. Johnstone and R. Thorpe, *Immunochromatry in Practice* (Blackwell, Oxford, 1982), pp. 48-56.
19. We thank K. E. Hellstrom and co-workers for help performing Fab assays, R. Kodori for the *pelB* gene cassette, R. Rahbar and B. Haselbeck for technical assistance, and R. Wall and G. Wilcox for comments on the manuscript. This work was supported by INGENE.

13 January 1988; accepted 18 March 1988

Technical Comments

Carcinogenic Risk Estimation

In their widely publicized and popularized article "Ranking possible carcinogenic hazard," Bruce N. Ames *et al.* (17 Apr. 1987, p. 271) conclude that "analysis on the levels of synthetic pollutants in drinking water and of synthetic pesticide residues in foods suggests that this pollution is likely to be a minimal carcinogenic hazard relative to the background of natural carcinogens" and thus that the "high costs of regulation" of such environmental carcinogens are unwarranted. These conclusions reflect both flawed science and public policy.

Although Ames *et al.* challenge the validity of animal carcinogenicity data for quantitative estimation of human risk, they nevertheless use such extrapolations, based on the percentage Human Exposure dose/Rodent Potency dose (HERP), for ranking carcinogenic hazards. Apart from the fact that HERP rankings are based on average population exposures excluding sensitive subgroups, such as pregnant women, the derived potencies of Ames *et al.*, doses inducing tumors in half the tumor-free animals, are misleading. Potencies for "synthetic pol-

lutants," such as trichloroethylene, are derived from bioassays in which lowest doses are large fractions of the maximally tolerated dose (MTD), whereas potencies for more extensively studied "natural carcinogens," such as aflatoxins, are generally derived from titrated doses, orders of magnitude below the MTD. Since dose-response curves are usually flattened near the MTD (1), potencies derived from high-dose testing yield artificially low risk estimates; HERPs for "synthetic" carcinogens are thus substantially underestimated compared with many "natural carcinogens."

Compounding this misconception, Ames *et al.* maintain that carcinogenic dose-response curves rise more steeply than linear curves and that tumor incidences increase more rapidly than proportional to dose. At high doses, dose-response curves are usually less steep than linear curves (1), as also recognized elsewhere by Ames and his colleagues (2). Thus at MTD doses, large further dose increases may induce only small increases in tumor incidence, perhaps reflecting competition between transforma-

tion and cytotoxicity (3); linear extrapolations from high-dose tests thus underestimate low-dose risks.

For Ames *et al.*, the term "carcinogen" heterogeneously includes direct and indirect influences, including promoting and modifying factors and mutagens. Caloric intake is considered "the most striking rodent carcinogen." However, no correlations have been established between food intake and tumor incidence among animals eating ad libitum, despite wide variations in caloric intake and body weight (4), nor have correlations been established between obesity and most human cancers. In the statement by Ames *et al.*, "at the MTD a high percentage of all chemicals might be classified as 'carcinogens'," toxicity and carcinogenicity are confused. However, among some 150 industrial chemicals selected as likely carcinogens and tested neonatally at MTD levels, fewer than 10% were carcinogenic (5). Many highly toxic chemicals are noncarcinogenic, and carcinogen doses in excess of the MTD often inhibit tumor yields. While Ames *et al.* revive the discredited theory that chronic irritation causes cancer, most irritants are noncarcinogenic, and there is no correlation between nonspecific cell injury and carcinogenic potency (6).

Ames *et al.* classify ethanol as carcinogenic, "[one of the two] largest identified causes

THIS PAGE BLANK (USPTO)

AMERICAN
ASSOCIATION FOR THE
ADVANCEMENT OF
SCIENCE

FAL

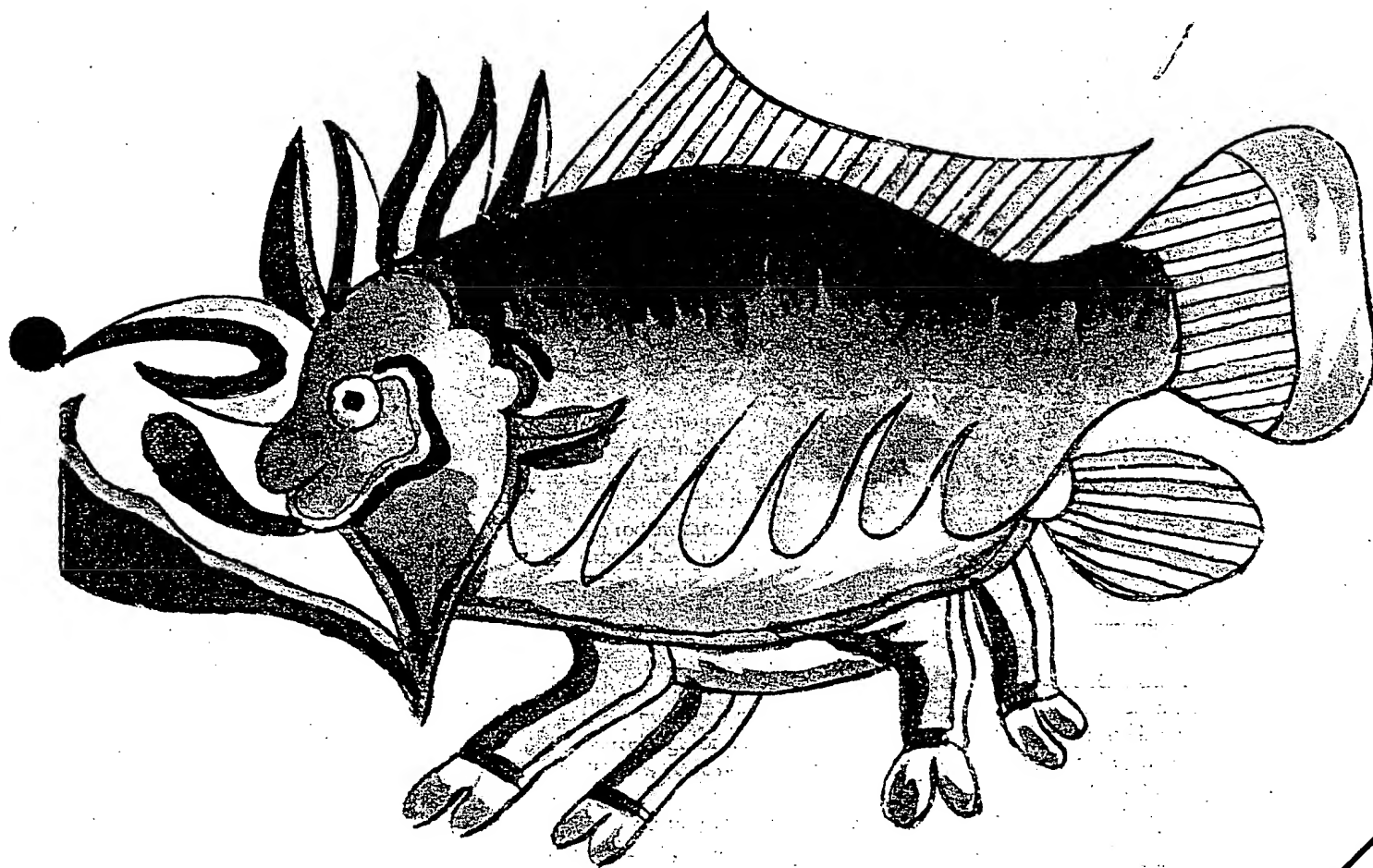
SCIENCE

FALCONER BIOL. LIB.

20 MAY 1988

\$3.00

VOL. 240 ■ PAGES 957-III2



BOOK ISSUE

THIS PAGE BLANK (USPTO)

fusion occurs at one cell membrane. Further, recent studies on glucose-mediated insulin release²¹ and the effects of CO₂ on vasopressin-induced water permeation²² suggest that cell acidification may be a more widespread stimulus for exocytosis than previously thought.

This work was supported by NIH grants AM 20999 and NS 07190 and by the New York Heart Association. J.v.A. is a Fellow of the National Kidney Foundation of New York and New Jersey.

Received 6 November 1984; accepted 18 January 1985.

1. Tycko, B. & Maxfield, F. R. *Cell* 28, 643-651 (1982).
2. Gluck, S., Cannon, C. & Al-Awqati, Q. *Proc. natn. Acad. Sci. U.S.A.* 79, 4327-4331 (1982).
3. Gluck, S., Kelly, S. & Al-Awqati, Q. *J. biol. Chem.* 257, 9230-9233 (1982).
4. Forgas, M., Canley, L., Wiedemann, B., Altstiel, L. & Branton, D. *Proc. natn. Acad. Sci. U.S.A.* 80, 1300-1303 (1983).
5. Stone, D. K., Xie, X.-S. & Racker, E. *J. biol. Chem.* 258, 4059-4062 (1983).
6. Galloway, C. J., Dean, G. E., Marsh, M., Rudnick, G. & Mellman, I. *Proc. natn. Acad. Sci. U.S.A.* 80, 3334-3338 (1983).
7. Glickman, J., Croen, K., Kelly, S. & Al-Awqati, Q. *J. Cell Biol.* 97, 1303-1308 (1983).
8. Rees-Jones, R. & Al-Awqati, Q. *Biochemistry* 23, 2336-2340 (1984).
9. Gluck, S. & Al-Awqati, Q. *J. clin. Invest.* 73, 1704-1710 (1984).
10. Schwartz, G. J. & Al-Awqati, Q. *J. clin. Invest.* (in the press).
11. Reeves, W., Gluck, S. & Al-Awqati, Q. *Kidney Int.* 23, 232 (Abstr.) (1983).
12. Steinmetz, P. R. *Physiol. Rev.* 54, 890-956 (1974).
13. Baker, P. F. & Honerjager, P. *Nature* 273, 160 (1978).
14. Lea, T. J. & Ashley, C. C. *Nature* 175, 236-238 (1978).
15. Lorenzen, M., Lee, C. O. & Windhager, E. E. *Fedn Proc.* 41, 1350 (1982).
16. Alvarez-Leefmans, F. J., Rink, T. J. & Tsein, R. Y. *J. Physiol., Lond.* 315, 531-548 (1981).
17. Tsien, R. Y. *Biochemistry* 19, 2396-2403 (1980).
18. Tsien, R. Y., Pozzan, T. & Rink, T. J. *J. Cell Biol.* 94, 325-334 (1982).
19. van Adelsberg, J. S. & Al-Awqati, Q. *Kidney Int.* 27, 290 (abstr.) (1985).
20. Roos, A. & Boron, W. *Physiol. Rev.* 61, 296-434 (1981).
21. Cook, D. L., Ikeuchi, M. & Fujimoto, W. Y. *Nature* 311, 269-271 (1984).
22. Lorenzen, M., Taylor, A. & Windhager, E. E. *Am. J. Physiol.* 245, F188-F197 (1983).
23. Rossier, M., Rossier, B. C., Pfeiffer, J. & Krachenhuhl, J. P. *J. Membrane Biol.* 48, 141-166 (1979).

The synthesis and *in vivo* assembly of functional antibodies in yeast

Clive R. Wood, Michael A. Boss, John H. Kenten,
Jane E. Calvert*, Nicola A. Roberts†
& J. Spencer Emtage

Celltech Ltd, 244-250 Bath Road, Slough, Berkshire SL1 4DY, UK

† Department of Biochemistry, University of Oxford,
South Parks Road, Oxford OX1 3QU, UK

The yeast *Saccharomyces cerevisiae* can synthesize, process and secrete higher eukaryotic proteins¹⁻⁵. We have investigated the expression of immunoglobulin chains in yeast and demonstrate here (1) the synthesis, processing and secretion of light and heavy chains, (2) the glycosylation of heavy chain, (3) the intracellular localization of these foreign proteins by immunofluorescence, and (4) the detection of functional antibodies in cells co-expressing both chains. This may provide the basis of a microbial fermentation process for the production of monoclonal antibodies. The co-expression of light and heavy chains in *Escherichia coli* has been reported but functional antibodies were not assembled *in vivo*^{6,7}. Furthermore, only low-level assembly of these chains was found *in vitro*.

The immunoglobulin λ and μ complementary DNAs used here were isolated from the mouse hybridomas S43 and B1-8, respectively^{8,9}. Both hybridomas were raised against the hapten 4-hydroxy-3-nitrophenyl acetyl (NP). The λ and μ cDNAs were placed under the control of the yeast 3-phosphoglycerate kinase (PGK) promoter, on pMA91 bearing the *LEU2* selectable marker¹⁰ (Fig. 1), to form pY λ for the expression of pre- λ , and pY μ for the expression of pre- μ . For co-expression of λ and μ on different plasmids in the same cell, another selectable marker was needed. pLG89 contains two selectable markers, *URA3* and *hph* (ref. 11), therefore the λ coding sequence and PGK promoter from pY λ were excised and inserted into pLG89, to produce pLGA.

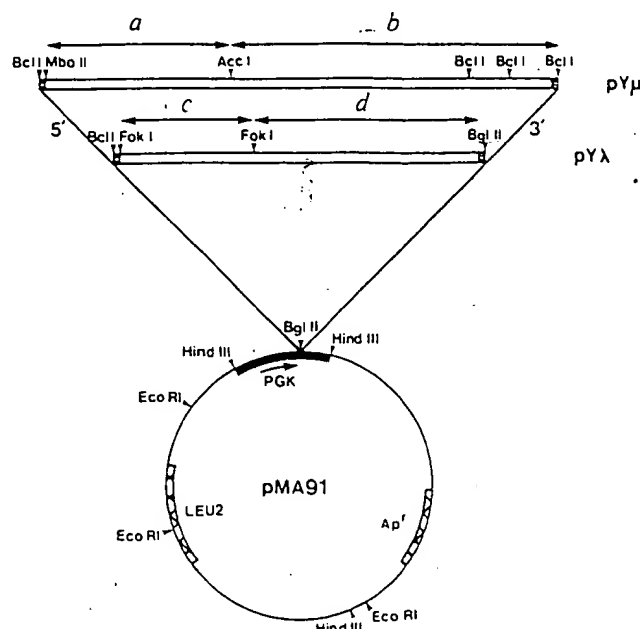


Fig. 1 Plasmids for the expression of immunoglobulin λ and μ chains in yeast. The figure shows the μ and λ inserts of pMA91 which were used to form pY μ and pY λ respectively. Solid bar, PGK sequences; the *LEU2* and *Ap'* marker genes are cross-hatched. Sequences derived from synthetic oligodeoxyribonucleotides are indicated by horizontal bars. An arrow indicates the direction of PGK transcription.

Methods. All DNA manipulations were carried out as described elsewhere²². For construction of pY λ , the *Hind*III fragment of pATA1-15 (ref. 6) bearing the λ cDNA was isolated, and cut with *Fok*I; 307-base pair (bp) *Fok*I and 512-bp *Fok*I-*Hind*III fragments were isolated (fragments c and d, respectively). Both oligodeoxyribonucleotides²³: R162 (5'-GATCAATGGCCTG-GATT 3') and R163 (5'-GTGAAATCCAGGCCATT 3') and pCT54 (ref. 24) cut with *Bcl*I and *Hind*III, to form pCTY λ . R162/R163 recreate the 5' coding sequence, and place a *Bcl*I site immediately upstream of the initiation ATG. pCTY λ was digested with *Hind*III, blunted with *S*₁ nuclease and filled-in with T4 DNA polymerase (P-L Biochemicals); the product was ligated with *Bgl*II linkers (5'-AGAGATCTCT 3'), then digested with *Bcl*I and *Bgl*II. The fragment bearing the λ cDNA was isolated and ligated with *Bgl*II-cut pMA91, to form pY λ . The 5' λ coding sequence of pY λ was isolated and shown to be correct by nucleotide sequencing. The λ coding sequence and PGK promoter were excised from pY λ on a *Hind*III fragment, and ligated with *Hind*III-cut pLG89 to form pLGA. For construction of pY μ , the μ cDNA was excised from pCT54 μ ²⁵ on a *Hind*III fragment, blunted (as for pCTY λ) and ligated with linker R107 (5'-TTTGTGATCAAAA 3') which contains an internal *Bcl*I site. The ligation product was digested with *Bcl*I and *Acc*I, and the fragment containing the 3' end of μ coding sequence isolated (b). Only the *Bcl*I site created by R107 will cut, for internal sites were *dam*-methylated. pCT54 μ was cut with *Mbo*II, ligated with R121 (5'-GATCAATGGGATGGAGCTGT 3') and R112 (5'-CAGCTCCATCCCAT 3') and digested with *Acc*I. The 276-bp *Bcl*I-*Acc*I fragment (fragment a) generated was isolated from a 5% polyacrylamide gel. pMA91 cut with *Bgl*II was ligated with both a and b, to form pY μ . The proximal end of the μ gene was isolated and shown to be correct by nucleotide sequencing^{26,27}. Yeast transformations were carried out as described elsewhere²⁸. pY λ and pY μ were individually transformed into MD46 (*a*/+ *pep4.3*/*pep4.3* *leu2*/*leu2* *arg5.6*/+ + *trp1* + *rad52* *ade1* + *his3*/+; Melanie Dobson, personal communication); pY μ and pLGA were individually transformed or co-transformed into X4003-5B (*a* *leu2* *ade1* *his4* *met2* *ura3* *trp5* *gal1*; Yeast Genetic Stock Center, Berkeley, California). Cells containing pY λ or pY μ were selected for by growth in the absence of L-leucine, and pLGA cells selected for by growth in the absence of uracil.

Cell extracts were subjected to Western blot analysis, using antisera to λ or μ proteins. MD46 cells containing pY λ were found to contain an immunoreactive protein (Fig. 2a, lane 3) that co-migrates with authentic B1-8 λ (Fig. 2a, lane 1) and with the mature λ protein synthesized by *E. coli* cells containing

* Present address: Department of Pathology, The Medical School, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne NE1 7RU, UK.

THIS PAGE BLANK (USPTO)

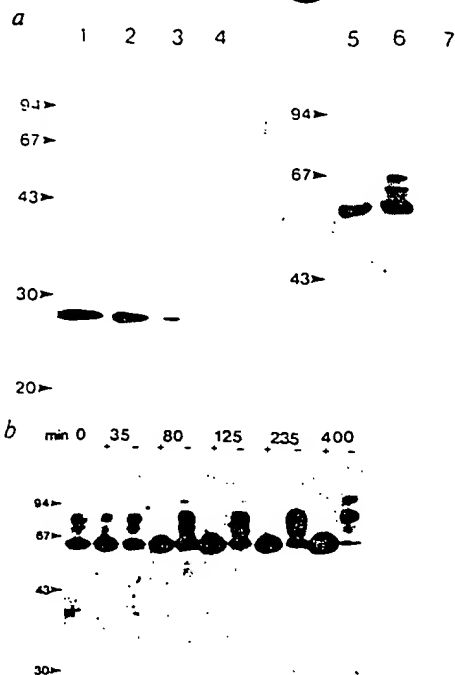


Fig. 2 Analyses of λ and μ proteins in yeast cell extracts. Cells were grown in YMM (2% (w/v) glucose, 0.67% (w/v) Difco yeast nitrogen base without amino acids) to $A_{660} = 1.0$, at 30°C with shaking, then collected by centrifugation. Cell pellets were disrupted by vortexing with glass beads (40 mesh, BDH) in 50 mM Tris-HCl pH 7.6, 1 mM EDTA, and diluted in sample buffer to 1.5% (w/v) SDS, 2.5% (v/v) β -mercaptoethanol, 5% (v/v) glycerol. Samples were subjected to SDS-polyacrylamide gel electrophoresis²⁹ and transferred to 0.45- μ m nitrocellulose and Western-blotted^{30,31} with either rabbit anti-mouse λ antibody (Miles) or rabbit anti-mouse IgM (Bionetics), and challenged with affinity-purified iodinated protein A ($2 \mu\text{Ci ml}^{-1}$; Amersham). Migration of yeast λ and μ was compared with that of B1-8 protein and of mature λ and μ synthesized in *E. coli*⁶. **a**, Lanes 1-4, samples blotted with anti- λ antibody: lane 1, B1-8; lane 2, bacterial λ ; lane 3, pY λ /MD46 cell extract; lane 4, MD46 cell extract. Lanes 5-7, samples blotted with anti-IgM: lane 5, bacterial μ ; lane 6, pY μ /MD46 cell extract; lane 7, MD46 cell extract. Arrowheads indicate the positions of markers (Pharmacia): phosphorylase b, 94 K; bovine serum albumin, 67 K; ovalbumin, 43 K; carbonic anhydrase, 30 K; soybean trypsin inhibitor, 20.1 K. For secretion studies, cells were also grown in YPED (1% (w/v) Difco Bacto yeast extract, 2% (w/v) Difco Bacto peptone, 2% (w/v) glucose) and all cultures were buffered with potassium phosphate pH 7.0. In these conditions, we were able to detect the ADH activity in extracts from $<0.005 A_{660}$ units of cells, which would represent lysis of $<0.5\%$ of a culture at $A_{660} = 1.0$. At such concentrations, ADH activity remained stable for up to 3 h at 30°C. Medium supernatants were assayed for ADH by mixing in a cuvette 1 ml of supernatant with 0.3 ml of 500 mM potassium phosphate pH 7.0, 100 μ l 30 mM NAD (reduced form; Sigma), 0.3 ml 30 mM acetaldehyde (BDH) and 1.3 ml double-distilled water. The decrease in A_{340} was measured with a Gilford spectrophotometer at 30°C. **b**, Demonstration of μ glycosylation in yeast. Duplicate 50-ml YMM cultures of MD46 containing pY μ were grown in 250-ml baffled flasks, with shaking at 30°C to $A_{660} = 1.0$ then split into duplicate 20-ml cultures. To one culture, designated '+', we added tunicamycin^{12,32} to a concentration of $15 \mu\text{g ml}^{-1}$, in 30 μ l of dimethyl sulphoxide (DMSO), and to the other culture, designated '-', we added 30 μ l of DMSO alone. Cultures were left shaking at 30°C, and samples taken at various times from both cultures. Cell extracts were prepared and Western-blotted with rabbit anti-mouse IgM and iodinated protein A. Alternating + and - tunicamycin samples are shown and the times of sampling given above the lanes in min. The 0-min sample was taken immediately before DMSO addition. Cells harbouring pY λ were treated similarly, and no difference was found between cultures with or without tunicamycin, when Western-blotted with anti- λ antiserum (data not shown). This result was expected as λ is a non-glycosylated protein in mammalian cells.

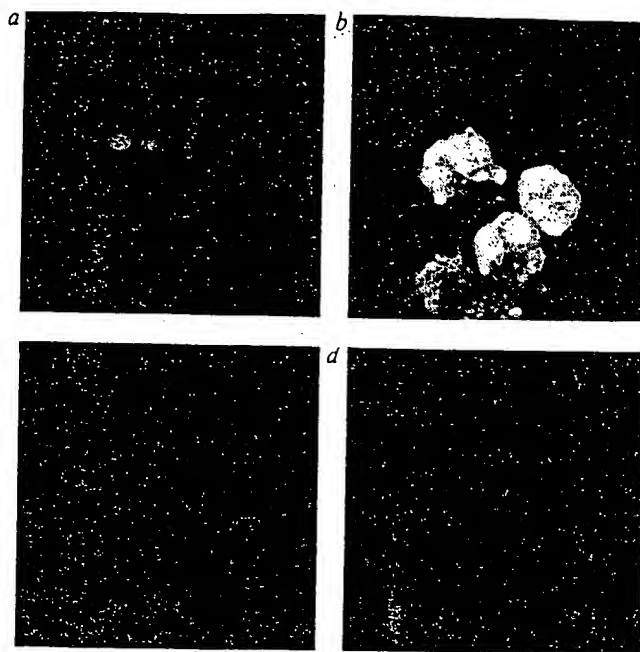


Fig. 3 Intracellular localization of λ and μ protein. The figure shows MD46 cells either untransformed (**c**, **d**), or transformed with pY μ (**a**) or pY λ (**b**); **a** and **c** were incubated with fluorescein-conjugated anti- μ ; **b** and **d** with rhodamine-conjugated anti- λ . $\times 700$.

Methods. Cells were grown in minimal media and converted to spheroplasts with zymolyase 20,000 (50 μg per A_{660} unit in 0.5 ml; Miles), in 1.2 M sorbitol, 50 mM potassium phosphate pH 7.0, 15 mM β -mercaptoethanol, 10 mM EDTA, at 30°C for 30 min. Spheroplasts were washed in 1.2 M sorbitol, and deposited on glass microscope slides using a cytocentrifuge, then fixed in 5% (v/v) acetic acid/95% (v/v) ethanol, at -20°C overnight. Slides were rehydrated and washed thoroughly in phosphate-buffered saline (PBS), before being stained with 0.1 mg ml^{-1} of fluorescein isothiocyanate-labelled affinity-purified goat anti-mouse IgM or tetramethylrhodamine isothiocyanate-labelled affinity-purified goat anti-mouse λ (both from Southern Biotechnology Associates, Alabama). Slides were washed in PBS, and mounted in glycerol containing 1,4-diazabicyclo(2.2.2)octane, before examination by ultraviolet microscopy.

a gene for mature λ (ref. 6) (Fig. 2a, lane 2). MD46 cells containing pY μ were found to contain three immunoreactive species (Fig. 2a, lane 6). The predominant yeast μ -immunoreactive band had a relative molecular mass (M_r) of $\sim 63,500$ (63.5 K) consistent with it being mature, non-glycosylated μ , and comigrated with the mature μ synthesized by *E. coli* cells containing a gene for mature μ (ref. 6) (Fig. 2a, lane 5). In addition, two diffuse bands of greater relative molecular mass were found in MD46 cells containing pY μ (Fig. 2a, lane 6). The mature, glycosylated μ of B1-8 has a $M_r \sim 70$ K (not shown).

We conclude that the λ immunoreactive species is mature λ , and the ~ 63.5 K μ species is mature, non-glycosylated μ . On the basis of relative molecular mass, the signal sequences of both proteins have probably been cleaved. The higher- M_r μ bands were shown by two criteria to be glycosylated μ . First, the higher M_r μ was lost from cells treated with tunicamycin, an inhibitor of *N*-linked glycosylation¹². In cultures without tunicamycin, the amount of higher- M_r μ increased during the time course examined, up to 235-min sample, after which the culture reached stationary phase, and the level of μ decreased (Fig. 2b). In cultures containing tunicamycin, the 63.5 K μ species accumulated with this species alone being found after 80 min of tunicamycin treatment. In addition, the higher- M_r form of μ was reduced in size to that of mature, non-glycosylated μ from *E. coli*, when treated with trifluoromethane sulphonic acid, which removes *N*-linked glycosyl groups with high efficiency¹³ (data not shown). We conclude that a significant proportion of yeast μ is *N*-glycosylated. However, the yeast

THIS PAGE BLANK (USPTO)

and mammalian μ will probably differ in carbohydrate composition, especially if the yeast μ has outer-chain oligosaccharides^{14,15}.

Medium supernatants from cultures of MD46 cells transformed with either pY λ or pY μ were found to contain immunoreactive λ or μ protein, respectively. This was shown to be genuinely secreted material, rather than arising from cell lysis, by demonstrating the absence of a cytoplasmic enzyme, alcohol dehydrogenase (ADH), from the supernatants (see Fig. 2 legend). In yeast minimal medium (YMM), up to 10% of λ and 5% of μ in MD46 cultures at $A_{660} \sim 1.0$ was found in the medium supernatant, as detected by enzyme-linked immunosorbent assay (ELISA)²⁵. In YPED medium (Fig. 2 legend) up to 40% of λ and 15% of μ was found in this fraction. Transformed X4003-5B cells generally yielded higher levels of secreted material. The immunoreactive λ material from the medium supernatants of cultures containing pY λ was shown, by Western blotting, to co-migrate with intracellular λ (data not shown). The secreted μ has not yet been examined.

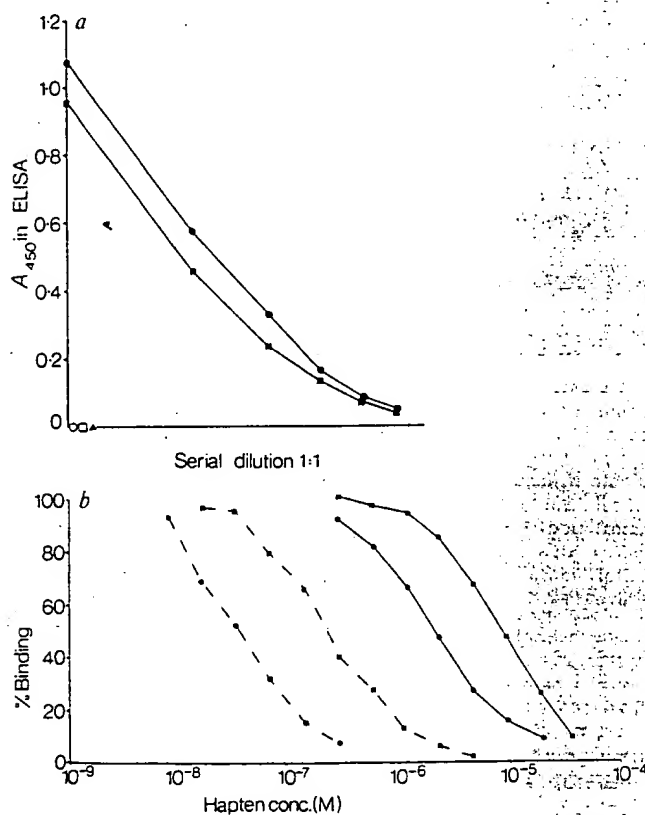


Fig. 4 Analysis of functional antibodies in yeast. Transformed X4003-5B cells were grown in 100-ml volumes of YMM, supplemented with 50 mg l⁻¹ of L-tryptophan, L-histidine and L-methionine, and 30 mg l⁻¹ of adenine sulphate in 500-ml flasks. Strains not containing pLG λ were supplemented with 30 mg l⁻¹ uracil, and those not containing pY μ with 50 mg l⁻¹ L-leucine. Cells were collected by centrifugation, then washed and resuspended in 25 mM Tris-HCl pH 8.0, 1 mM EDTA, 5% (v/v) glycerol, 0.1% (v/v) Nonidet-P40, 1 mM phenylmethylsulphonylfluoride (extraction buffer), before being lysed by glass bead disruption. 'Insoluble' material was pelleted in a microfuge for 12 min at 4 °C, and the 'soluble' supernatant fraction removed. The soluble fraction was analysed in a two-site sandwich ELISA (NIP assay) which detects μ -chain binding to haptenylated bovine serum albumin (NIP-caproate-BSA)⁶; this assay is sensitive to 60 pg of B1-8 IgM. *a*, Specific hapten binding was examined in the NIP assay of the soluble fraction of pLG λ + pY μ -transformed cells (○), pY μ -transformed cells (△) and B1-8 (■, □), in the absence (solid symbols) and presence (open symbols) of 30 μ M free NIP-caproate. *b*, The heteroclitic nature of yeast antibody activity (●) in soluble fractions from pLG λ + pY μ -transformed cells and B1-8 (■) was examined by comparing binding in the presence of free NIP-caproate (---) and NP-caproate (—).

The intracellular locations of λ and μ proteins in yeast were examined by staining fixed yeast spheroplasts with fluorescein or rhodamine-conjugated antisera. We stained both MD46 and X4003-5B strains, harbouring pY μ , pY λ , pLG λ or pYMA. pYMA is a pMA91 derivative encoding the mature λ protein, that is, Met-mature λ , not including the signal sequence. Cells containing pY μ (Fig. 3a), pYMA (Fig. 3b), pY λ or pLG λ showed a discrete immunofluorescence localized in bodies that appeared on the basis of their morphology to be vacuoles^{16,17}. Cells containing pYMA showed a much greater accumulation of stain than those containing pY λ ; this may be explained by the observation that the intracellular concentration of mature λ from pYMA is up to fourfold greater than that of pre- λ from pY λ , in MD46, by ELISA. Further experiments are required to identify unequivocally the sites of accumulation of immunoreactive λ and μ proteins. If these foreign proteins are being localized in the vacuole, they could be transported there by the secretory pathway^{18,19}, however, the detection of mature λ from pYMA in the same structures calls this into question—the cell may be directing these proteins to the vacuole for degradation²⁰.

Soluble extracts of X4003-5B cells transformed with one or both of pY μ and pLG λ were prepared and analysed by an ELISA that detects binding of B1-8 to solid-phase hapten, in the presence and absence of competing free hapten⁶ (Fig. 4a); no specific binding was detected for extracts of cells transformed with either pLG λ or pY μ . However, extracts of cells transformed with both plasmids, and expressing both light and heavy chains, showed a strong, specific signal similar to that of the hybridoma B1-8 protein. All detectable solid-phase hapten binding was lost in the presence of 30 μ M free hapten. B1-8 is a heteroclitic antibody, and shows greater affinity for the related hapten 4-hydroxy-5-iodo-3-nitrophenyl acetyl (NIP), than for the hapten NP²¹. The yeast antibody activity also showed that higher free NP than free NIP concentrations were required to inhibit binding of antibody to solid-phase NIP (Fig. 4b). In addition, the activities of the yeast antibody and B1-8 showed similar specificity ratios (ratio of concentration of NP to NIP at 50% inhibition) of 37 and 47, respectively. The amount of μ protein in the soluble extract of cells containing pLG λ and pY μ was determined by ELISA, and using B1-8 as a standard, the specific activity of the yeast antibody was found to be ~0.5%. ELISA showed that at $A_{660} = 1.0$, there was ~500 ng of B1-8 λ equivalent, and 15 ng of B1-8 μ equivalent per ml. No significant amounts of NIP binding activity have been found in concentrated culture media (data not shown).

On the basis of the specific activity of the soluble-fraction yeast antibodies, the efficiency of assembly of functional antibodies is low, although the antibodies show both specific hapten binding and heterocliticity. It will be interesting to characterize further the yeast antibodies in the soluble fraction and to determine whether or not most of the immunoglobulin protein from the insoluble fraction (~75% of total) is also assembled into functional antibodies.

We thank Drs A. Bothwell and D. Baltimore for supplying plasmids pAB λ 1-15 and pAB μ -11; Drs M. Neuberger and T. Imanishi-Kari for gifts of B1-8 protein; Dr J. E. Davies for the provision of pLG89; and Drs M. Dobson, A. J. Kingsman and S. M. Kingsman (Oxford) for pMA91 and MD46. We thank our many colleagues in Celltech and Oxford, for advice and discussions.

Received 8 October, 1984; accepted 22 January 1985.

1. Hitzeman, R. A. *et al. Science* **219**, 620-625 (1983).
2. Rothenstein, S. J., Lazarus, C. M., Smith, W. E., Baulcombe, D. C. & Gatenby, A. A. *Nature* **308**, 662-665 (1984).
3. Valianzuola, P., Medina, A., Rutter, W. J., Ammerer, G. & Hall, B. D. *Nature* **298**, 347-350 (1982).
4. Hitzeman, R. A. *et al. Nucleic Acids Res.* **11**, 2745-2763 (1983).
5. Brake, A. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4642-4646 (1984).
6. Boss, M. A., Kenten, J. H., Wood, C. R. & Emtage, J. S. *Nucleic Acids Res.* **12**, 3791-3806 (1984).
7. Cabilly, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3273-3277 (1984).
8. Bothwell, A. L. M. *et al. Cell* **24**, 625-637 (1981).
9. Bothwell, A. L. M. *et al. Nature* **298**, 380-382 (1982).
10. Mellor, J. *et al. Gene* **24**, 1-14 (1983).
11. Griz, L. & Davies, J. *Gene* **25**, 179-188 (1983).

THIS PAGE BLANK (USPTO)

12. Mahoney, W. C. & Duskin, D. *J. biol. Chem.* **254**, 6572-6576 (1979).
13. Edge, A. S. B., Fallynck, C. R., Hof, L., Reichert, L. E. & Weber, P. *Analyt. Biochem.* **118**, 131-137 (1981).
14. Chapman, A. & Kornfeld, R. *J. biol. Chem.* **254**, 816-823 (1979).
15. Ballou, C. E. in *The Molecular Biology of the Yeast Saccharomyces, Metabolism and Gene Expression* (eds Strathern, J. et al.) 335-360 (Cold Spring Harbor Laboratory, New York, 1982).
16. Schwencke, J. *Physiol. Vég.* **15**, 491 (1977).
17. Maillie, P. A. *Rev. Pl. Physiol.* **29**, 193 (1978).
18. Schekman, R. & Novick, P. in *The Molecular Biology of the Yeast Saccharomyces, Metabolism and Gene Expression* (eds Strathern, J. et al.) 361-398 (Cold Spring Harbor Laboratory, New York, 1982).
19. Stevens, T., Esmon, B. & Schekman, R. *Cell* **30**, 439-448 (1982).
20. Lin, C.-J., Chopra, A. K., Srmadova, M. & Chaloupka, J. *FEMS Microbiol. Lett.* **21**, 313-317 (1984).
21. Imanishi, T. & Makela, O. *Eur. J. Immun.* **3**, 323-330 (1973).
22. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
23. Patel, T. P. et al. *Nucleic Acids Res.* **10**, 5605-5620 (1982).
24. Emtage, J. S. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3671-3675 (1983).
25. Wood, C. R., Boss, M. A., Patel, T. P. & Emtage, J. S. *Nucleic Acids Res.* **12**, 3937-3950 (1984).
26. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
27. Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 302-321 (1981).
28. Beggs, J. *Nature* **275**, 104-109 (1978).
29. Lacemli, U. K. *Nature* **227**, 680-685 (1970).
30. Burnette, W. N. *Analyt. Biochem.* **112**, 195-203 (1981).
31. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
32. Onishi, H. R., Tkacz, J. S. & Lampen, J. O. *J. biol. Chem.* **254**, 11943-11952 (1979).

Functional modifications of cytotoxic T-lymphocyte T200 glycoprotein recognized by monoclonal antibodies

Leo Lefrançois & Michael J. Bevan

Department of Immunology, Scripps Clinic and Research Foundation, 10666 North Torrey Pines Road, La Jolla, California 92037, USA

Plasma membrane glycoproteins of cytotoxic T lymphocytes (CTLs) are involved in the binding to and subsequent destruction of appropriate target cells¹⁻³. The electrophoretic profile of surface proteins of mature CTLs, particularly those of high relative molecular mass (M_r), is markedly different from that of naive peripheral T cells or non-cytolytic T cells⁴⁻⁷, suggesting the possible involvement of these molecules in the activation of CTLs and/or in the lytic process itself. By generating monoclonal antibodies to cell-surface proteins of CTL clones, we have now detected CTL-specific modifications in one of these high- M_r membrane proteins, T200. Although forms of T200 are found on a wide variety of cell types, the neoantigenic determinants recognized by our antibodies are present exclusively on activated T cells and in high concentrations only on CTLs. Furthermore, the expression of the modifications recognized by our antibodies is influenced by soluble factors and also seems to have functional significance, as monoclonal antibodies specific for these novel epitopes block cytolytic activity.

Monoclonal antibodies with specificity for CTL surface recognition structures were produced by immunizing BALB.B mice repeatedly (intraperitoneally) with a C57BL/6-derived CTL clone, B3.3, which is specific for a BALB minor histocompatibility antigen in association with H-2K^b. Spleen cells of immune mice were fused to a myeloma partner, P3-X63 Ag8.653, and the resulting hybridoma supernatants were screened for the ability to block specific target lysis by B3.3. Figure 1a depicts the blocking activity of two of these monoclonal antibodies, CT1 and CT2, on clone B3.3 in conditions of saturating antibody. CT1 and CT2 (which are both IgM antibodies) efficiently block specific killing of BALB.B target cells by clone B3.3, CT2 being the more efficient of the two antibodies. In this regard, CT2 was as efficient as the anti-Lyt-2 monoclonal antibody, 53.6.72. Antibody 13/2.3, which is specific for the T200 molecule⁸, did not significantly affect specific lysis by this CTL clone.

To determine whether CT1 and CT2 were able to block CTL-mediated killing in conditions not requiring antigen-specific recognition on the part of the CTLs, we tested their ability to inhibit lectin-dependent cell-mediated cytotoxicity (LDCC). Both antibodies were highly efficient as inhibitors of

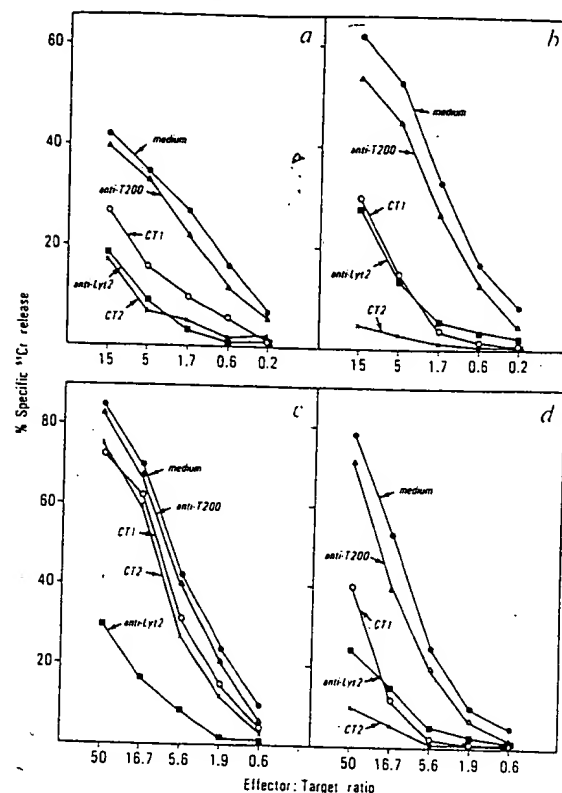


Fig. 1 Inhibition by CT antibodies of cytotoxicity mediated by cloned CTLs and MLC-derived CTLs. Serial dilutions of effector cells were incubated in 96-well round-bottomed microtitre plates for 30 min at 4°C with saturating levels of CT1 (○), CT2 (×), 13/2.3 (Δ, anti-T200, kindly provided by Dr Ian Trowbridge, Salk Institute), 53.6.72 (■, anti-Lyt-2), or medium (●). Effector cells were: a, b, CTL clone B3.3 (a C57BL/6 clone derived by limiting dilution and specific for a BALB minor histocompatibility antigen in association with H-2K^b); c, C57BL/6 spleen cells stimulated for 5 days with irradiated DBA/2 spleen cells (5×10^6 cells of each per well in 24-well plates) in RPMI 1640 culture medium supplemented with 5% fetal calf serum; d, MLC cells derived as in c but restimulated (5×10^5 responders and 5×10^6 stimulators per well) every 7 days for 4 weeks in medium containing 5% rat Con A supernatant (RCS). ⁵¹Cr-labelled target cells (1×10^4) were as follows: a, 3-day BALB.B Con A blasts; b, P815 H-2^d mastocytoma cells with the addition of $10 \mu\text{g ml}^{-1}$ PHA; c, d, P815 cells. After the addition of target cells, the plates were centrifuged for 3 min at 800 r.p.m. to initiate cell contact and incubated at 37°C for 2 h. Per cent specific lysis was calculated as $100 \times [(c.p.m. \text{ released with effectors}) - (c.p.m. \text{ released alone})] / [(c.p.m. \text{ released by detergent}) - (c.p.m. \text{ released alone})]$. Spontaneous release of P815 target cells was 5% and that of BALB.B Con A blasts 18%. No significant lysis was observed of C57BL/6 Con A blasts (a), of P815 cells without the addition of PHA (b), or of EL4 cells (c, d).

phytohaemagglutinin (PHA)-dependent killing of P815 tumour cells by clone B3.3 (Fig. 1b). CT1 was comparable with anti-Lyt-2 in blocking LDCC, while the anti-T200 antibody did not inhibit lysis (Fig. 1b). (Note that CT1 and CT2 do not bind to the target cells used in these assays; see below.) We also found that five out of five independently isolated CTL clones of various target specificities were blocked by both antibodies. Thus, it was apparent that CT1 and CT2 were not clone-specific or anti-idiotypic in their reactivity. More surprising results were obtained when we examined their effect on heterogeneous populations of CTLs generated in mixed lymphocyte culture (MLC). When the effects of these antibodies on the lytic ability of a primary MLC were tested (Fig. 1c), only minimal inhibition of specific lysis was observed using CT1, CT2 or the control antibody 13/2.3 (anti-T200), while 53.6.72 (anti-Lyt-2) significantly inhibited lysis. However, quite surprisingly, in inhibition assays using as effectors CTLs from a long-term MLC, CT1 and CT2 produced a significant decrease in specific lysis (Fig. 1d). This MLC (and others similarly blocked by CT1 and CT2) were

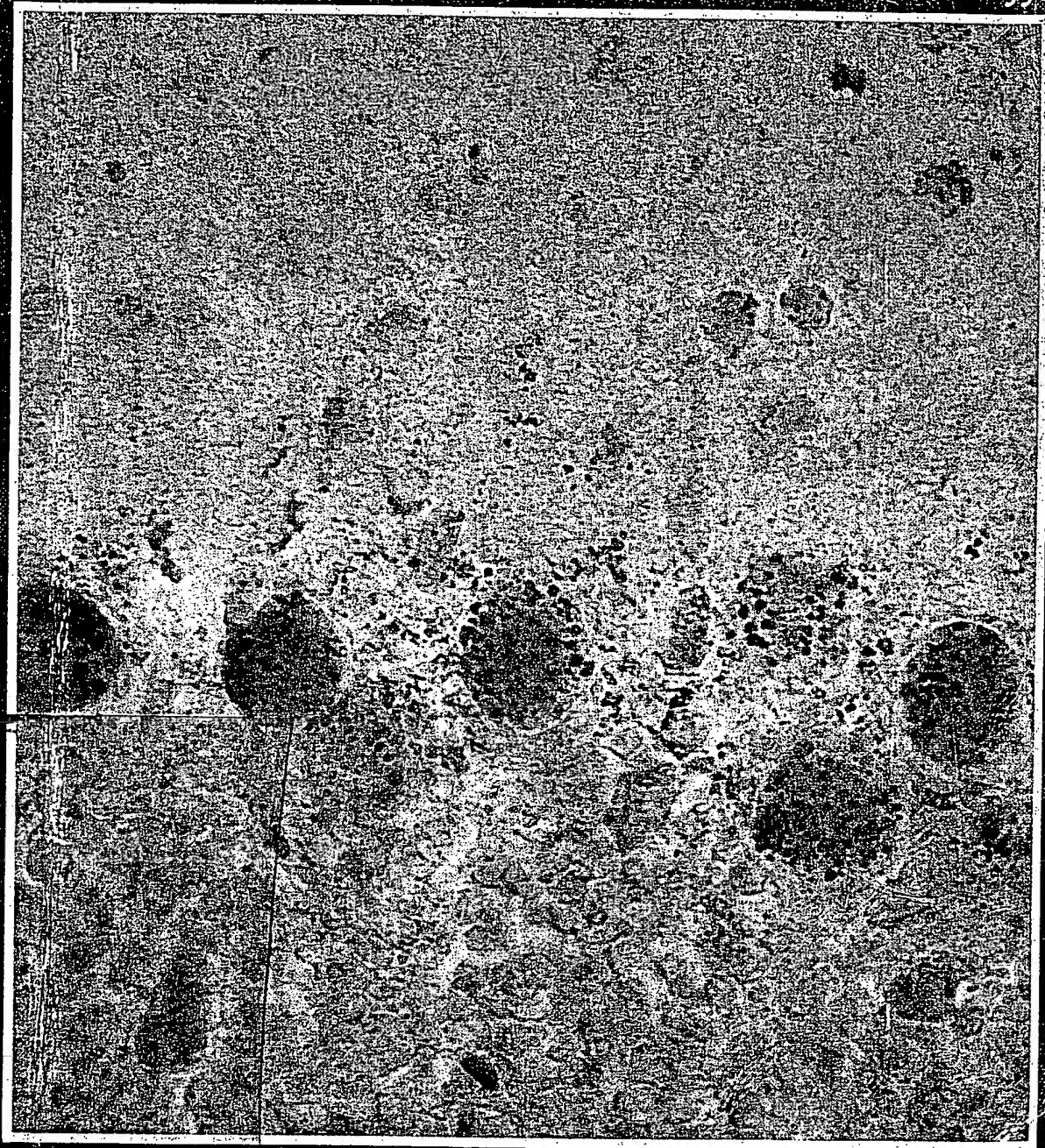
THIS PAGE BLANK (USPTO)



Cells & Organelles

THIRD EDITION

Holtzman & Novikoff



THIS PAGE BLANK (USPTO)

Address orders to:
383 Madison Avenue
New York, NY 10017

Address editorial correspondence to:
West Washington Square
Philadelphia, PA 19105

Text Typeface: Souvenir
Compositor: The Clarinda Company
Acquisitions Editor: Michael Brown
Project Editor: Sally Kusch
Copyeditor: Elizabeth Galbraith
Managing Editor & Art Director: Richard L. Moore
Art/Design Assistant: Virginia A. Bollard
Text Design: Caliber Design Planning, Inc.
Cover Design: Lawrence R. Didona
Text Artwork: J & R Technical Services, Inc.
Production Manager: Tim Frelick
Assistant Production Manager: Maureen Iannuzzi

Front cover credit: Rat cerebellum section. Immunocytochemical localization of lysosomal β -galactosidase. (*Journal of Cell Biology*, in press.) Micrograph by Phyllis M. Novikoff.
Back cover credit: Rat hepatocytes surrounding pancreatic islet cells. Red spheres indicate cytosolic lipid. (*Proc. Natl. Acad. Sci. USA*, 1979.) Micrograph by Phyllis M. Novikoff.

Library of Congress Cataloging in Publication Data

Holtzman, Eric, 1939-
Cells and organelles.

Rev. ed. of: Cells and organelles/Alex B.
Novikoff, Eric Holtzman. 2nd ed. c1976
Includes bibliographies and index.

1. Cytology. 2. Cell organelles. I. Novikoff, Alex
Benjamin, 1913- II. Novikoff, Alex
Benjamin, 1913- Cells and organelles. 2nd
ed. III. Title. [DNLM: 1. Cells. 2. Organoids-QH
581.2 N943c]

QH581.2.H64 1983 574.8'7 83-7683

ISBN 0-03-049461-3

Cells and Organelles (3/e)

ISBN 0-03-049-461-3

© 1984 by CBS College Publishing. Copyright 1970-1976 by Holt, Rinehart and Winston.
All rights reserved. Printed in the United States of America.
Library of Congress catalog card number 83-7683.

3456 032 987654321

CBS COLLEGE PUBLISHING
Saunders College Publishing
Holt, Rinehart and Winston
The Dryden Press

THIS PAGE BLANK (USPTO)

one another can fall into rhythms on the basis of effects of these sorts, without special coordinating devices. With cilia this could account for the fact that experimental changes in the viscosity (resistance to flow) of the medium in which ciliates are placed can markedly alter the pattern of ciliary waves. Hydrodynamic interactions could also explain findings made on the flagellate *Myxotrichia*, which possesses, in addition to its flagella, a symbiotic population of spirochetes (Section 3.2.4) associated with its surface. Despite the absence of an obvious coordinating device such as a common plasma membrane or interconnected basal structure, these spirochetes, whose intrinsic motion is somewhat akin to that of a flagellum, beat in waves. It even appears that motion of the attached spirochetes can produce motion of the protozoan itself.

3.3.5 Endosymbionts

Many protozoa maintain long-term or permanent symbiotic relationships with other cells. *Myxotrichia* with its spirochetes was just mentioned. Particularly interesting from the viewpoint of cellular evolution are the *endosymbiotic relationships* in which one cell lives inside another. Entry most often is by endocytic mechanisms; the entrant is presumably protected by mechanisms such as the ones discussed in Section 2.8.2. Both blue-green algae and eucaryotic algae have been found thriving inside species of protozoa, apparently "contributing" photosynthetic products and O_2 to the host and "receiving" the respiratory product CO_2 for use in photosynthesis, among other benefits. Bacteria also live inside protozoa. In one classic example, bacteria were originally detected in *Paramecium* through findings that some *Paramecium* strains could release particles capable of killing other strains. The particles turned out to be small bacteria normally harbored in vacuoles within the "killer" strains. Another case is the amoeba *Pelomyxa*, in which no recognizable mitochondria are present but in which there are populations of endosymbiotic bacteria perhaps capable of providing the amoeba with respiratory capacities.

As with other symbioses, the relative balances of benefits between the partners vary considerably, and the situation can approach parasitism. In many cases, one or the other partner can survive without the symbiotic association. However, a majority of students of cellular evolution believe that certain endosymbiotic relationships have evolved into virtually absolute interdependencies and that of these, some ancient ones account for the origin of intracellular organelles, especially chloroplasts and mitochondria. This belief will be pursued further in Section 4.5.2.

Chapter 3.4 Eucaryotic Plant Cells

Until recently, discussions of eucaryotic plant cells were frequently introduced with explanations about the slow progress of plant cell biology. Plants do pose

THIS PAGE BLANK (USPTO)

special problems in preservation for microscopy and in biochemical analysis owing in large measure to their walls, which, for example, complicate the homogenization steps for cell fractionation. Many of these problems have been overcome, however, and progress in plant cell biology is accelerating rapidly. This is important both for the basic scientific information being gained and for practical matters of agriculture and ecology.*

Algae other than the blue-greens and all higher plants are eucaryotes. With very few exceptions their cells contain the organelles and exhibit the metabolic capacities described in Part 2. Also described in Part 2 was one of the characteristic cell organelles limited to plants, the plastids (Chap. 2.7). However, in addition to their photosynthetic capacities, plants show important differences from animals that extend to the cellular level: For example, the vascular system of higher plants carries water, minerals absorbed by the roots, and carbohydrates being transported from sites of synthesis in green tissues to sites of use or storage. A variety of small molecules are also translocated, including hormones. Unlike in higher animals, the extracellular fluids of the vascular system are not rich in macromolecules, nor is there an abundant population of circulating cells. Paralleling this, plant cells are surrounded by walls that permit passage of gases, water, inorganic ions, and other small molecules into and out of the cells but severely restrict access of larger molecules to the cell surface and also limit the interactions of adjacent cells with one another. As with bacteria, the walls provide mechanical support so that plants can grow in fresh water, and multicellular forms can circulate a relatively dilute solution; the cells can resist osmotically induced pressures without bursting. These characteristics, plus the presence in the plant cell of vacuoles controlling ion distributions within the cell, produce balances of ion and water movements that can be quite different from the ones usually found in animal cells.

What do these factors imply for the plant cell plasma membrane? It seems likely that the receptors, ion pumps, and other plasma membrane molecules are somewhat different from those in animals. There is fragmentary information suggesting that this is the case—for instance, as will be seen in subsequent sections, plant scientists have stressed transport of K^+ and H^+ across the plasma membrane more than the Na^+ transport of such strong interest to investigators of animal cells (Section 2.1.3). However, even such major relevant issues as the roles of cyclic nucleotides in plant cells are still quite unsettled. One question of interest is whether plants possess plasma membrane receptors that bind macromolecules and whether they engage in much endocytosis. At first glance neither endocytosis nor macromolecular binding would seem likely to be a major normal event for the plant cell surface given the restrictions imposed by the wall. But this probably is an oversimplified view. Excellent photomicrographs have been published showing coated vesicles apparently budding from the cell surface of algae, and when cell walls are exper-

*Readers with a special interest in plant science should refer to the index listing "Plant cell(s)" for material covered in this text, and to the Further Reading section at the end of this part.

imer
cells
surfa
Lecti
(Sect

Nitro
and p
lular r
patho
major
negati
growth
ivities
and by
the bac
there v
Among
plant ce
that ger
roots. O
establish
vices, sc
lectins (S
der suspi
One

rived from
quired for
much rich
sources. (I
peas, or c
nitrogen.)
ductive to
hemoglobi
of the cells
oxygen the
terial meta
natives of
which the f
between ho

In add
to certain p
specialized f
duced to ni
tase). The ni

THIS PAGE BLANK (USPTO)

imentally removed and macromolecular tracers are introduced, some plant cells seem able to take up the tracers. At the very least this might reflect cell surface turnover phenomena of the sorts discussed in Sections 2.1.6 and 2.8.4. Lectins capable of recognizing and binding specific carbohydrate configurations (Section 2.1.7) also are present at plant cell surfaces.

Nitrogen Fixation Another kind of evidence for endocytosis-like capacities, and perhaps for cell surface recognition systems that cooperate with intracellular receptors, comes from the responses of plant cells to invasive organisms—pathogens, parasites, and symbionts. For instance, the root nodules found on major crop plants like peas and beans are sites of interactions between Gram-negative bacteria of the genus *Rhizobium* and the plants' root cells. Through growth of infective threadlike structures that penetrate the roots, enzymatic activities and production of hormone-like signaling molecules by the plant host and by the invading bacteria, a sequence of events is orchestrated whereby the bacteria come eventually to reside in the host cell cytoplasm. They live there within cell surface invaginations and membrane-bounded vacuoles. Among other processes, the bacterial entry requires localized dismantling of the plant cell wall and induces a pattern of growth and division of the host cells that generates the characteristic lumplike bacteria-containing nodules in the roots. Only certain species of host plants and specific strains of *Rhizobium* can establish the relationships described; this specificity argues for recognition devices, some of which may reside in the cells' walls or membranes. (Host cell lectins [Section 2.1.7] recognizing bacterial cell surface polysaccharides are under suspicion.)

Once inside, the *Rhizobium* cells ("bacteroids") use carbohydrates derived from the host plant's metabolism to support the energy production required for the fixation of atmospheric nitrogen. The plant is provided with a much richer nitrogen supply than that available from the nitrates and other soil sources. (This, of course, is also the basis for crop rotation in which soybeans, peas, or other legumes are planted periodically to re-enrich the soil with usable nitrogen.) The host cell apparently aids in maintaining an environment conducive to the functioning of the bacteroids it contains. A protein related to hemoglobin [leghemoglobin] is synthesized by the host. Along with the walls of the cells in the nodules, this substance limits and controls the amount of free oxygen that reaches the bacteria, permitting adequate amounts to support bacterial metabolism without inactivating nitrogen fixation (Chap. 3.2B). The derivatives of the host cells' plasma membranes, which surround the spaces in which the *Rhizobium* cells live, probably participate actively in the exchanges between host and bacterium.

In addition to the symbiosis-like relations just outlined, which apply only to certain plants, the nitrogen metabolism of plants in general involves other specialized features. For example, nitrates absorbed from the soil are first reduced to nitrites by a cytoplasmic (cytosolic?) enzyme system (nitrate reductase). The nitrites are then converted to ammonia and its derivatives by nitrite

THIS PAGE BLANK (USPTO)

reductases and other systems in the chloroplasts that use the photosynthetic apparatus as the source of reducing capacities. Peroxisomal uric acid oxidase (Section 2.9.2) may also take part in the metabolic conversions by which nitrogen is made available for amino acid synthesis and other pathways.

Chapter 3.4A Algae

Algae are plants—some unicellular (such as *Chlamydomonas*; Figs. II-59 and II-61), others multicellular (some such as the giant kelps are quite large). They are classified partly on the basis of their color, which is determined largely by the nature of the pigments present in addition to the chlorophyll. For example, in red algae a major additional pigment is similar to the red pigment of blue-green algae (Chap. 3.2C).

As Figure III-17 indicates, the chloroplast membranes of algae are usually less elaborately developed than in higher plants (see also Fig. II-59). The plastids are membrane-bounded, and the internal sacs or thylakoids are arranged as long parallel sheets, with several adjacent thylakoids often associating more or less closely for substantial distances. (See Section 2.7.1. These groups are sometimes called "lamellae.") Most algae lack the stacked-coin-like grana characteristic of higher plants (see Figs. II-59 and II-60). The plastids contain ribosomes and fibrils of DNA. A *pyrenoid*, where starch is stored and probably synthesized, and a *pyrenoid sac*, which also stores polysaccharides, are part of the plastid.

The mitochondria may be like those in Figure III-17, or the cristae may be tubular as in many protozoa. The ER is relatively scanty; as usual, it is continuous with the nuclear envelope. The Golgi apparatus often lies near the nucleus in algae, and it seems likely that membrane from the nuclear envelope feeds into the outer surface of the Golgi stack through vesicles, as portrayed in the diagram. Lysosomes and peroxisomes have not been isolated from algae, but both organelles have been reported by electron microscopists. Autophagic vacuoles have been seen to increase in number in starved *Euglena*, the chloroplast-containing protist closely related to algae. In *Euglena*, too, the number of peroxisomes has been shown to increase greatly when the algae are shifted from a glucose medium to one with acetate or alcohol, suggesting to the investigators that the peroxisomes may function in lipid and carbohydrate metabolism as they do in seeds (Section 2.9.2). In contrast to most higher plant cells, paired centrioles are present in many algae.

Many algae have prominent cell walls based on cellulose, as in higher plants, but some do not. The wall surrounding *Chlamydomonas*, for example, is rich in glycoproteins and lacks cellulose; it more resembles an animal cell's coat than a typical wall of a higher plant. In certain algae the cell walls show specialized plaquelike scales. Small vacuoles, derived from the Golgi apparatus, transport the scales and other material to the forming cell walls. In diatoms, the walls are made largely of silica.



Figure III-17
Not included in cytoplasmic granules extended parallel are referred to as well. (Compare

With a large is possible to su vacuole that occ of the concepio These cells also ity; their size fac central vacuole. of microfilament myosin was usec Some unic with the cell, cc 2.7.4). We have,

THIS PAGE BLANK (USPTO)

Synthetic Signal Peptide and Analogs Display Different Activities in Mammalian and Plant *in Vitro* Secretion Systems*

(Received for publication, September 2, 1986)

Le T. Duong†, Michael P. Caulfield, and Michael Rosenblatt

From the Parathyroid Hormone Laboratory, Biological Research and Molecular Biology Division, Merck Sharp and Dohme Research Laboratories, West Point, Pennsylvania 19486

The biological properties of four chemically synthesized signal peptides were compared in mammalian (rabbit reticulocyte) and plant (wheat germ) cell-free protein secretion systems. The precursor-specific region of bovine pre-parathyroid hormone (preproPTH), [D-Tyr-(+1)]preproPTH-(−29+1)amide, and a sulfur-free analog, [Nle-(−25), Nle-(−21), Nle-(−18), Ala-(−14), D-Tyr-(+1)]preproPTH-(−29+1)amide, inhibit the processing of an unrelated precursor protein (pre-prolactin) to its mature secreted form (prolactin) in the mammalian system. In the plant system supplemented with signal recognition particle, the signal peptides arrest translation of both secretory (pre-prolactin) and cytoplasmic (globin) proteins. One analog, [Nle-(−25), Nle-(−21), Asp-(−18), Ala-(−14), D-Tyr-(+1)]preproPTH-(−29+1)amide, inhibits pre-protein processing in the mammalian system but fails to induce translation arrest in the plant system. A truncated peptide, [N^α-AcLeu-(−17), Ala-(−14), D-Tyr-(+1)]preproPTH-(−17+1)amide, lacking the N-terminal (positively charged) region and a portion of the hydrophobic core region, is inactive in both systems. These studies demonstrate that the chemically synthesized signal region of a precursor protein interacts directly with signal recognition particle and functionally mimics the proposed properties of a native signal sequence linked to a nascent protein as it emerges from the ribosome during biosynthesis, and an analog of the signal peptide reveals fundamental differences between the components involved in the protein secretion apparatus in mammals and plants.

A number of secretory preproteins, lysosomal proteins, and integral plasma membrane proteins have been shown to interact with the signal recognition particle (SRP)¹ in order to be correctly translocated across the rough endoplasmic retic-

ular (RER) membrane (1, 2). SRP has been postulated to possess at least three discrete activities critically related to the physiological secretion of proteins by cells. First, SRP is thought to bind the signal sequence of nascent proteins as it emerges from the ribosome (3, 4). Second, this interaction of a signal sequence with SRP leads to the formation of an SRP-ribosome complex which specifically inhibits further translation of precursor protein (5). Third, SRP is recognized by an SRP receptor (6–8) or docking protein (9) in the RER membrane. This latter interaction presumably initiates and orients the attachment of the ribosomes to the RER membrane, establishing the required spatial arrangement for protein co-translational translocation across the membrane and reversing the translational arrest which was in effect (5–9).

The phenomenon of SRP-mediated translation arrest has so far only been detected *in vitro* in a heterologous system employing plant (wheat germ) ribosomes and mammalian (canine) SRP. In homologous mammalian systems, such as the rabbit reticulocyte lysate or HeLa cell-free systems, no such SRP arrest of protein elongation has been observed (10). However, SRP appears to be endogenous to mammalian lysates and causes translational arrest of precursor proteins when added to the wheat germ system (9). Therefore, it is unclear what role SRP plays in the translation of secretory proteins in the mammalian system. Although the SRP-mediated translation arrest is not required for protein translocation across the RER *in vitro* (11), it has been suggested that it occurs physiologically, thus preventing completion of biosynthesis of secretory proteins in the cytoplasm (2).

In order to study the biological properties of the signal region of a precursor protein and its interactions with components of the mammalian and plant cellular secretory apparatus, we chemically synthesized the precursor-specific region of bovine pre-parathyroid hormone, [D-Tyr-(+1)]preproPTH-(−29+1)amide, and three analogs of this sequence (Table I).

In the reticulocyte lysate, the native-sequence signal peptide was previously demonstrated to competitively inhibit the translocation and processing of bovine pre-parathyroid hormone, bovine pre-prolactin, and human pre-placental lactogen (12). The finding that a synthetic signal peptide of one protein could inhibit the translocation and processing of unrelated precursor proteins, derived from differing organs and species (i.e. that it competes for recognition of a component(s) of the eukaryotic secretory apparatus), indicated that the synthetic signal peptide shares common conformational features and functionality with other native signal sequences. Others also reported a similar observation using a consensus signal peptide in the same assay system (13). However, it remains to be determined at what level and with what specific component(s) of the secretory apparatus the synthetic signal peptide exerts its effect. The synthetic signal peptides could

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom reprint requests should be addressed.

The abbreviations used are: SRP, signal recognition particle; RER, rough endoplasmic reticulum; RM, canine pancreatic rough microsomal membranes; K-RM, salt-washed RM; pre-Prl, preprolactin; Prl, prolactin; PTH, parathyroid hormone; native-sequence peptide, [D-Tyr-(+1)]preproPTH-(−29+1)amide; sulfur-free peptide, [Nle-(−25), Nle-(−21), Nle-(−18), Ala-(−14), D-Tyr-(+1)]preproPTH-(−29+1); truncated peptide, [N^α-AcLeu-(−17), Ala-(−14), D-Tyr-(+1)]preproPTH-(−17+1)amide; Asp-substituted peptide, [Nle-(−25), Nle-(−21), Asp-(−18), Ala-(−14), D-Tyr-(+1)]preproPTH-(−29+1)amide; Nle, norleucine; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; preproPTH, pre-parathyroid hormone.

THIS PAGE BLANK (USPTO)

interact with either SRP or another membrane-associated component in the mammalian system.

Since the activity of SRP can only be assayed in the wheat germ system, we examined the properties of the synthetic native-sequence signal peptide and its analogs in this system. In addition, we compared the biological properties of these signal peptides in the plant system to the mammalian system. Our studies indicate that synthetic signal peptides can functionally mimic the biological properties of native signal regions present in nascent precursor proteins as they emerge from the ribosomes and that they can interact directly with SRP, even when unattached to a nascent precursor protein. Furthermore, the biological properties of one of the synthetic peptide analogs reveals differences in the nature of the mammalian *versus* the plant protein secretory apparatus. These results are discussed in relation to the current working hypothesis (1, 2) for the mechanism of cellular protein secretion.

EXPERIMENTAL PROCEDURES

In Vitro Translation/Processing Assay—Rabbit reticulocyte lysate (New England Nuclear) and wheat germ lysate (Bethesda Research Laboratories) cell-free translation systems were used as described by the suppliers. [³⁵S]Methionine (New England Nuclear) was used at 2.7 mCi/ml. Bovine pituitary mRNA was extracted (14) and translated *in vitro* at a concentration of 1.5 µg/ml and 15 µg/ml in the reticulocyte and wheat germ lysates, respectively. A mixture of rabbit α- and β-globin mRNAs (Bethesda Research Laboratories) was used at 0.6 µg/ml in the wheat germ system. SRP was purified from canine pancreatic microsomes using a ω-aminopentyl-agarose column and a sucrose gradient according to the procedure of Walter and Blobel (15). In this study, purified SRP (9.5 µg/ml) was added to the wheat germ translation system in a nonsynchronized fashion, since 7-methylguanosine 5'-phosphate was not included after 2 min of protein synthesis (5). Canine pancreatic rough microsomal membranes (RM) and salt-washed RM (K-RM) were prepared as described (16). All membrane preparations were treated with staphylococcal nuclease (Boehringer Mannheim). RM were used at 0.3 and 0.5 A₂₆₀ units/ml in the reticulocyte and wheat germ lysates, respectively. K-RM were used at 0.5 A₂₆₀ units/ml in both lysates. The peptide concentration used in the bioassays was calculated from the peptide content determined by amino acid analysis.

The sequence of addition of K-RM and synthetic signal peptides to the reaction mixtures was found to be critical in the wheat germ system but not in the reticulocyte system. Unless otherwise stated, synthetic signal peptide and/or SRP were added first to the wheat germ lysate and were incubated for 5 min on ice. K-RM were then added. Protein synthesis was initiated by addition of mRNAs and incubation at the appropriate temperature. Translations were performed for 1 h at 37 °C for the reticulocyte system and at 25 °C for the wheat germ system. Reactions were stopped by the addition of 1% sodium dodecyl sulfate (SDS) and boiling for 5 min.

Synthesis and Purification of Synthetic Signal Peptides—[D-Tyr(-1)]preproPTH(-29→+1)amide (native-sequence peptide) was synthesized and purified as previously described (17). [Nle(-25), Nle(-21), Nle(-18), Ala(-14), D-Tyr(+1)]preproPTH(-29→+1)amide (sulfur-free peptide) was synthesized on an Applied Biosystems

430A automated peptide synthesizer and purified on a Vydac C-4 high performance liquid chromatography column using a 35–41% aqueous acetonitrile (containing 0.1% trifluoroacetic acid) gradient. The sulfur-free peptide was rerun on the same column and eluted isocratically as a single peak at 37% acetonitrile. A truncated peptide lacking the positively charged N terminus and a portion of the hydrophobic region, [N^α-AcLeu(-17), Ala(-14), D-Tyr(+1)]preproPTH(-17→+1)amide (truncated peptide) was isolated as a minor product from the synthesis of the sulfur-free peptide. The truncated peptide was purified using a Vydac C-4 column and a 28–31% acetonitrile gradient. It was then further purified using an isocratic (28% acetonitrile) elution. An analog of the sulfur-free peptide, [Nle(-25), Nle(-21), Asp(-18), Ala(-14), D-Tyr(+1)]preproPTH(-29→+1)amide (Asp-substituted peptide), contains a negatively charged aspartic acid substitution within the hydrophobic region of the signal sequence. This peptide was synthesized in a similar manner to that described for the sulfur-free peptide. It was purified on a Vydac C-4 high performance liquid chromatography column with a 29–36% acetonitrile gradient and with an isocratic (36% acetonitrile) elution.

For each peptide, amino acid composition and sequencing analysis was performed. Amino acid sequences indicated a peptide purity of >95% for all peptides. The truncated peptide could not be directly examined by sequence analysis since it had a blocked N terminus. However, amino acid analysis of each tryptic fragment of this peptide indicated the presence of the expected sequence. Furthermore, we had evidence that the N terminus of this peptide was acetylated, and the molecular mass determined by mass spectroscopy agreed with that predicted for the acetylated species.

Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)—Total proteins from cell-free translation/processing assays were separated by SDS-PAGE and visualized by autoradiography. SDS-PAGE was performed as described (18) using 10–20% polyacrylamide gradient gels. After fixing, gels were treated with the autoradiographic enhancer Enlightening (New England Nuclear), dried, and exposed to SB-5 x-ray film (Eastman Kodak) at -80 °C. Amount of protein synthesis was roughly quantitated by determining the density of each protein band on the x-ray film using a Hoefer GS 300 Scanning Densitometer.

RESULTS

Activity of the Synthetic Signal Peptides in the Wheat Germ and the Rabbit Reticulocyte Systems—The amino acid sequences of all the synthetic signal peptides used in this study are shown in Table I. Earlier reports from this laboratory (12) have demonstrated that the native-sequence signal peptide inhibits the conversion of pre-prolactin (pre-Prl) to prolactin (Prl) and causes an accumulation of pre-Prl in the rabbit reticulocyte lysate supplemented with RM. In this study, the sulfur-free analog was demonstrated to display the same activity at a comparable concentration as the native-sequence peptide (Fig. 1A). At a concentration of 3 µM, the sulfur-free peptide does not affect the biosynthesis of pre-Prl (lanes 1 and 2) but inhibits the conversion of pre-Prl to Prl (lanes 3 and 4) in the presence of RM. This inhibition of processing

TABLE I
Amino acid sequences of the synthetic signal peptide and analogs of bovine pre-proparathyroid hormones

Abbreviated Name	Amino Acid Sequence									
	-29	-25	-20	-15	-10	-5	-1	+1		
native-sequence ⁽¹⁾	H ₂ N-SerAlaLysAspMetValLysValMetIleValMetLeuAlaIleCysPheLeuAlaArgSerAspGlyLysSerValLysArg								(D)-TyrCONH ₂	
sulfur-free ⁽²⁾	H ₂ N	Nle	Nle	Nle	Ala					(D)-TyrCONH ₂
truncated ⁽³⁾			Ac-HN	Ala						(D)-TyrCONH ₂
Asp-substituted ⁽⁴⁾	H ₂ N	Nle	Nle	Asp	Ala					(D)-TyrCONH ₂

¹ [D-Tyr(+1)]preproPTH(-29→+1) amide.

² [Nle(-25), Nle(-21), Nle(-18), Ala(-14), D-Tyr(+1)]preproPTH(-29→+1) amide.

³ [N^α-AcLeu(-17), Ala(-14), D-Tyr(+1)]preproPTH(-17→+1) amide.

⁴ [Nle(-25), Nle(-21), Asp(-18), Ala(-14), D-Tyr(+1)]preproPTH(-29→+1) amide.

THIS PAGE BLANK (USPTO)

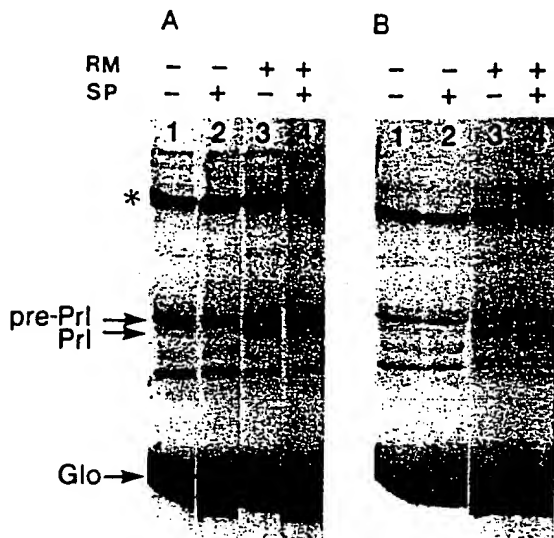


FIG. 1. Synthetic signal peptide of bovine pre-proparathyroid hormone inhibits processing of pre-prolactin in the rabbit reticulocyte lysate. Cell-free translation of bovine pituitary mRNA (1.5 $\mu\text{g}/\text{ml}$) in rabbit reticulocyte lysate was carried out as described under "Experimental Procedures." Total translation products were separated on SDS-PAGE and visualized by autoradiography. Arrows indicate the positions of pre-prolactin (pre-Prl) and prolactin (Prl). The presence (+) or absence (–) of various components added to the reaction mixture is indicated above each lane. A, biosynthesis of pre-Prl was in the absence (lane 1) or presence (lane 2) of 3 μM sulfur-free signal peptide (SP). In the presence of RM (0.3 A_{280} units/ml), processing of pre-Prl to Prl was examined without (lane 3) and with (lane 4) 3 μM of sulfur-free peptide. Positions of globin (Glo) and a higher molecular weight protein (*), which bound [^{35}S]methionine specifically, are also marked. B, translation and processing of pre-Prl were as described in A, except that 10 μM truncated signal peptide was added in lanes 2 and 4.

was demonstrated to be signal peptide-specific, since the truncated peptide, which lacks the N-terminal positive charge and a portion of the hydrophobic region, failed to inhibit processing of pre-Prl to Prl at a concentration as high as 10 μM (Fig. 1B, lanes 1–4). Other unrelated peptides, such as glucagon and fragments of PTH (bPTH-(1–34) and hPTH-(28–54)) also failed to inhibit processing of pre-Prl to Prl (data not shown).

In contrast to the reticulocyte system, the native-sequence and the sulfur-free peptides did not inhibit processing of precursor protein to its mature form in the wheat germ translation system supplemented with identical RM. As shown in Fig. 2A, 9 μM of native-sequence signal peptide did not cause an accumulation of pre-Prl (lane 4).

Synthetic Signal Peptides Cause an Irreversible SRP-dependent Translation Arrest—Purified SRP was assayed in the wheat germ translation system in the presence or absence of synthetic signal peptides. In the absence of signal peptides, SRP caused a specific translational arrest of pre-Prl, a secretory protein, but had no effect on the biosynthesis of globin, a nonsecretory protein (Fig. 2A, lane 5; Fig. 2B, lane 2). Since canine SRP was added in the translation mixture in a non-synchronized fashion, it did not completely inhibit the biosynthesis of pre-Prl in this experiment. However, this translation arrest was reversed when K-RM were added back to lysate (Fig. 2A, lane 7).

When either 9 μM of the native-sequence peptide (Fig. 2A, lane 6) or 6.7 μM of the sulfur-free peptide (Fig. 2B, lane 3) was added to the wheat germ in the presence of SRP, synthesis of both pre-Prl and globin was completely inhibited. As described under "Experimental Procedures," we have found that

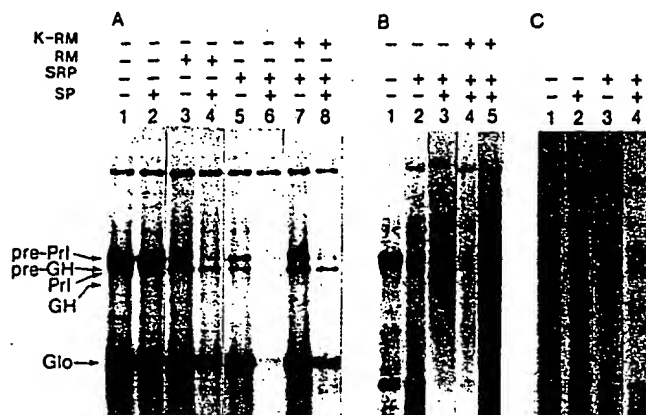


FIG. 2. Synthetic signal peptide inhibits translation of globin in the presence of signal recognition particle in the wheat germ lysate. Bovine pituitary mRNA (15 $\mu\text{g}/\text{ml}$) and globin mRNA (0.6 $\mu\text{g}/\text{ml}$) were translated in the wheat germ lysate, as described under "Experimental Procedures." Unless otherwise stated, signal peptide and SRP (9.5 $\mu\text{g}/\text{ml}$) were added to the translation mixture, on ice, and incubated for 5 min. RM (0.5 A_{280} unit/ml) or K-RM (0.5 A_{280} unit/ml) then were added. Protein synthesis was initiated by addition of mRNAs and incubated at 25 °C. Arrows indicate the following translation products: pre-Prl, Prl, pre-growth hormone (pre-GH), growth hormone (GH), and globin (Glo). The presence (+) or absence (–) of signal peptide (SP), SRP, and K-RM are indicated above each lane. A, native-sequence signal peptide (9 μM) was used in this experiment. Translation was performed without (lane 1) or with (lane 2) signal peptide. In the presence of RM, translation was without (lane 3) or with (lane 4) signal peptide. SRP was added to the translation mixture without (lane 5) and with (lane 6) signal peptide. Lanes 7 and 8 are the same as lanes 5 and 6, respectively, except that K-RM was included prior to the addition of signal peptide. B, sulfur-free signal peptide (6.7 μM) was used in this experiment. Translations were carried out without (lane 1) or with SRP (lane 2) or with SRP and signal peptide (lane 3). Lanes 4 and 5 are identical to translation mixture described in lane 3, except that K-RM were added at 5 min (lane 4) or 0 min (lane 5) after initiation of protein biosynthesis. C, truncated signal peptide (30 μM) was used in this experiment as a control. Translation of pituitary mRNA and globin mRNA in the wheat germ lysate in the absence (lanes 1 and 2) or presence (lanes 3 and 4) of SRP. Signal peptide was included in lanes 2 and 4.

the inhibition of protein synthesis was observed when SRP and the signal peptide were preincubated, even on ice, with the wheat germ lysate prior to the addition of mRNAs. If K-RM were added to the SRP-containing lysate prior to the addition of the sulfur-free signal peptides, protein synthesis was not inhibited (Fig. 2A, lane 8). Similarly, if K-RM were added at time 0 (Fig. 2B, lane 5), 30 sec, 1 min, or 2 min after initiation of protein synthesis at 25 °C, the synthesis of Prl and globin was normal. However, addition of K-RM 5 min after initiation of protein synthesis could no longer reverse the inhibition of Prl and globin synthesis (Fig. 2B, lane 4). The specific effect of the native-sequence and sulfur-free signal peptides in the wheat germ system was demonstrated by the failure of the truncated peptide (30 μM) to inhibit globin synthesis (Fig. 2C, lane 4).

The Synthetic Signal Peptide/SRP Translation Arrest Is Dependent on the Length of the Nascent Chain—To study the relationship of peptide elongation and the translation arrest activity of the synthetic signal peptide, sulfur-free signal peptide (6.7 μM) was added to the SRP-containing wheat germ lysate at time 0, 30 sec, 1, 2, 5, 10, and 20 min after initiation of globin synthesis (Fig. 3, lanes 2–8, respectively). As compared to the control amount of globin synthesis in the absence of the signal peptide (Fig. 3, lane 1), the addition of the signal peptide at 0, 30 sec, or 1 min after initiation of translation could inhibit approximately 90% of globin synthe-

THIS PAGE BLANK (USPTO)

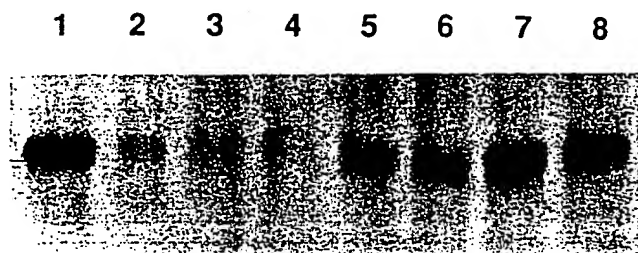


FIG. 3. Effect of peptide elongation on the translation arrest activity of the synthetic signal peptide in the wheat germ lysate. Globin mRNA (0.6 $\mu\text{g}/\text{ml}$) was translated in the presence of SRP (9.5 $\mu\text{g}/\text{ml}$) in the wheat germ lysate as described under "Experimental Procedures." Total translation products were separated on SDS-PAGE and visualized by autoradiography. Lane 1, control globin synthesis in the absence of signal peptide. Sulfur-free signal peptide (6.7 μM) was added to the translation mixture at time 0, 30 sec, 1, 2, 5, 10, and 20 min (lanes 2-8, respectively) after the initiation of synthesis.

sis (Fig. 3, lanes 2-4). After 2-5 min of peptide elongation, the signal peptide inhibited only ~40-60% of globin synthesis (Fig. 3, lanes 5 and 6). After 10-20 min of peptide elongation, the signal peptide had virtually no effect on globin translation (Fig. 3, lanes 7 and 8).

Synthetic Signal Peptides Do Not Cause an Inhibition of Protein Translation in the Rabbit Reticulocyte Lysate, Even in the Presence of Exogenously Added SRP—Other workers have demonstrated the presence of endogenous SRP-like activity in the rabbit reticulocyte lysate and the failure of exogenously added SRP to arrest translation of secretory proteins in this system (9). The finding of an irreversible inhibition of translation of pre-Prl and globin by signal peptides plus SRP in the wheat germ system prompted us to examine the effects of these peptides in the reticulocyte system supplemented with canine SRP. Fig. 4 (lanes 1 and 2) demonstrates that exogenously added SRP had no effect upon biosynthesis of pre-Prl in the absence or presence of sulfur-free peptide (6.7 μM). Furthermore, the same concentration of this peptide still inhibited processing of pre-Prl to Prl in the presence of canine SRP and K-RM (Fig. 4, lane 4). Similar results were obtained with the native-sequence signal peptide (data not shown).

A Negatively Charged Amino Acid in the Hydrophobic Core of the Signal Peptide Causes Loss of Recognition by SRP, but Retains the Property of Inhibiting Processing—An Asp-substituted analog has a substitution of aspartic acid for methionine at position (-)18 in the center of the hydrophobic region of the signal sequence (Table I). In the reticulocyte lysate system, this analog was demonstrated to retain signal peptide activity; namely, the peptide inhibited precursor protein processing and caused the accumulation of pre-Prl (Fig. 5, lanes 5 and 6), and this peptide had no effect on protein biosynthesis (Fig. 5, lanes 2 and 3). The Asp-substituted analog is approximately one-half as potent in inhibiting protein processing as the native-sequence or the sulfur-free peptides, since 5.2 μM of the Asp-substituted analog (Fig. 5, lane 6) was required to obtain a comparable degree of inhibition in pre-Prl processing as 3 μM of the sulfur-free peptide (Fig. 2A, lane 4). However, it is important to note that formal comparisons of potencies have not been generated for each of the peptides used in this study.

When the Asp-substituted analog was assayed in the wheat germ system in the presence of SRP, there was no effect on globin synthesis (Fig. 5, lanes 11 and 12; 5.2 and 7.8 μM of peptide, respectively). We have titrated the Asp-substituted signal peptide to 26 μM with no detectable translation arrest activity in the wheat germ system (data not shown).

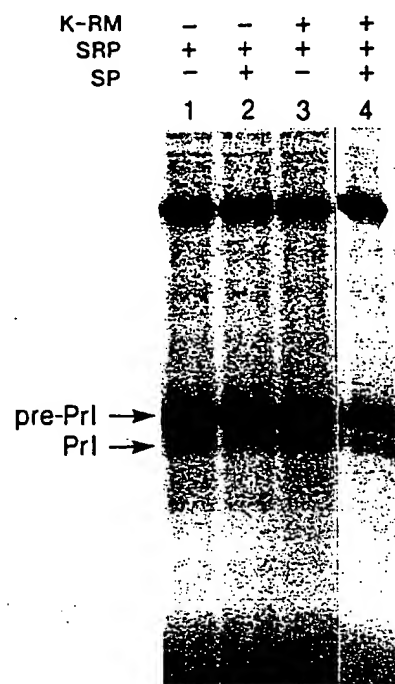


FIG. 4. Effect of synthetic signal peptide in the reticulocyte lysate supplemented with signal recognition particle. Bovine pituitary mRNA was translated in the reticulocyte lysate as described in Fig. 1. In the presence of added canine SRP (9.5 $\mu\text{g}/\text{ml}$), translations were carried out without (lane 1) or with (lane 2) sulfur-free signal peptide (SP) (6.7 μM). In the presence of K-RM, translations were without (lane 3) and with (lane 4) signal peptide. The presence (+) or absence (-) of individual components is indicated above each lane.

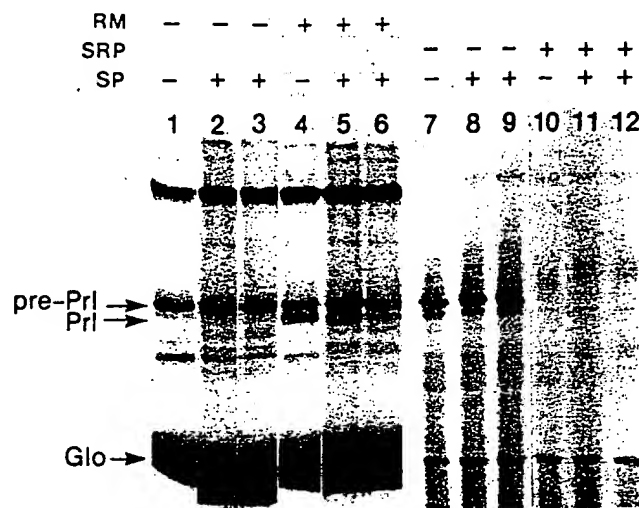


FIG. 5. Asp-substituted signal peptide inhibits processing of preprolactin in reticulocyte lysate but fails to inhibit globin synthesis in wheat germ system. Translation of pituitary mRNA in the reticulocyte lysate was carried out, as described in Fig. 1, in the absence (lanes 1-3) and presence (lanes 4-6) of RM. Asp-substituted signal peptide was included at 2.6 μM (lanes 2 and 5) or 5.2 μM (lanes 3 and 6). Translation of pituitary mRNA and globin mRNA in the wheat germ lysate was as described in Fig. 2 in the absence (lanes 7-9) and presence (lanes 10-12) of SRP. Asp-substituted signal peptide was included at 5.2 μM (lanes 8 and 11) or 7.8 μM (lanes 9 and 12). The presence (+) or absence (-) of various components added to the reaction mixture is indicated above each lane. SP, signal peptide; Glo, globin.

THIS PAGE BLANK (USPTO)

DISCUSSION

The chemical synthesis of the signal sequence of parathyroid hormone enables a direct examination of the biological properties of this specialized sequence when it is not contiguous with the remainder of the nascent precursor protein. The generation of analogs of this signal peptide permits structure-activity studies for various events in protein secretion in which a signal peptide is thought to play an important physiological role. Furthermore, in this study, such analogs were used to explore the nature of the differences between the mammalian and plant secretory apparatus.

The biological properties of the native-sequence synthetic signal peptide of PTH and its sulfur-free analog in the reticulocyte lysate (their ability to inhibit the conversion of pre-Prl to its mature form) demonstrate that synthetic signal peptides compete with a native precursor protein for specific sites or components of the secretory apparatus. These observations were made using PTH-related peptides (12) and using a consensus signal sequence (13) in the rabbit reticulocyte lysate translation/processing system. A likely and plausible site of action of the signal peptides in this system is the RER membrane. If this were the case, then inhibition of processing should also be seen when RER membranes are utilized with a different translation system. However, when these peptides were added to a plant (wheat germ) translation/processing system, in the presence of the same canine microsomal membranes, no inhibition of processing was observed. It therefore seems unlikely that the synthetic signal peptide is inhibiting processing by interacting directly with the microsomal membrane.

Since the synthetic signal peptides apparently do not interact with the membrane, they might interact with a soluble component which was required for protein secretion prior to the membranous site. A likely candidate for this component would be SRP. In the wheat germ system, SRP was demonstrated to interact with the signal sequence of pre-Prl as it emerged from the ribosome and arrested the polysomes from further translating pre-Prl mRNA (3-5). In support of the direct binding of SRP to a signal sequence of a nascent protein, other workers have shown that the signal sequence of pre-Prl can be cross-linked to the 54-kDa SRP subunit (19, 20). Thus, in this study we investigated the interaction of the synthetic signal peptides with SRP.

Addition of either the native-sequence synthetic signal peptide or its sulfur-free analog to the wheat germ translation system supplemented with SRP caused translation arrest not only of the secretory protein, pre-Prl, but also of the cytoplasmic protein, globin. The arrest of globin synthesis strongly suggests that the synthetic signal peptide can bind directly to SRP and induce the formation of a signal peptide-SRP-ribosome complex. This complex, in turn, produces a specific translation arrest even of mRNA for a cytoplasmic protein. This observation further suggests that the signal peptide-SRP-ribosome complex very closely resembles the complex formed as a native signal sequence (contiguous with a nascent protein) emerges from the ribosome. These data also indicate that the synthetic signal peptide adopts a conformation similar to a native signal sequence.

Unlike the SRP-induced inhibition of secretory protein biosynthesis, which can be reversed by the addition of K-RM (5), the synthetic signal peptide/SRP-dependent translation arrest cannot be readily reversed by the addition of K-RM. In order to prevent translation arrest, K-RM and SRP must be reconstituted in the wheat germ lysate prior to the formation of the synthetic signal peptide-SRP-ribosomal complex. The formation of this complex becomes irreversible after 5

min at 25 °C, since addition of K-RM at this time did not release the synthesis of pre-Prl and globin. Therefore, in the absence of K-RM, SRP associates with ribosomes in a configuration which favors a tight interaction with the synthetic signal peptide. In the presence of K-RM, SRP appears to be preferentially membrane-bound. It is not certain whether the membrane-bound SRP can still bind the synthetic signal peptide. Nevertheless, the inhibitory effect of this peptide on protein translation was diminished when SRP was reconstituted with K-RM.

The inhibitory effect of the signal peptide was also diminished if globin was translated for 10 min prior to the addition of the peptide. In our hands, the rate of peptide elongation is approximately 10 residues/min in the wheat germ translation system (data not shown). Therefore, nascent peptides of 100 residues were made in 10 min. The total lengths of α - and β -globin are 141 and 146 residues, respectively (21). Since 40 residues are assumed to be buried within the ribosome (5), 60 residues are thus exposed and effectively prevent the formation of the signal peptide-SRP-ribosomal complex. Since globin does not have a contiguous signal sequence, it is likely that the exposed 60 residues of globin sterically prevent the interaction of the synthetic signal peptide and/or SRP with the ribosomes.

The translation arrest activity of the synthetic signal peptide and SRP in the wheat germ system was investigated at 37 °C. However, no conclusion could be reached since the translation efficiency of the wheat germ lysate was significantly decreased at this higher temperature (data not shown). On the other hand, the addition of exogenous SRP and sulfur-free signal peptide to the reticulocyte lysate, which functions optimally at 37 °C, did not have any effect on protein synthesis.

Failure to see a translation arrest in the mammalian system in the presence of exogenously added SRP and synthetic signal peptides, conditions that cause a complete inhibition of protein synthesis in the plant system, indicates that SRP displays different properties in the two systems. This finding confirms a previous report (10) demonstrating that SRP-mediated translation arrest activity occurs in plant but not in mammalian, *in vitro* secretion systems. The observed change in SRP activity in these translation systems can be explained in at least two ways. First, SRP may interact differently with plant versus mammalian ribosomes, as suggested previously by Meyer (10). Second, the mammalian system may contain an additional factor(s) that interacts with the signal peptide, either in conjunction with or independently of SRP, and which is absent in the plant system. It has already been shown (10) that this factor(s), if present, is not a soluble docking protein-like molecule.

Another argument has been suggested that the rate of peptide elongation in the reticulocyte is much faster than that in the wheat germ system; therefore, the fast rate of synthesis prevents the interaction of SRP with the signal sequence as it emerges from the ribosomes (22). In this study, we have observed that a preincubation of the synthetic signal peptide and canine SRP in the reticulocyte lysate, either on ice or at 37 °C, prior to the addition of mRNA, did not inhibit the synthesis of pre-Prl. This suggested that the rate of synthesis was unlikely to affect the formation of signal sequence-SRP-ribosome complex.

The differences between mammalian and plant systems are highlighted by the profile of biological properties displayed by the Asp-substituted analog. This analog contains a kind of signal sequence modification studied previously in bacterial systems (23): the insertion of a negatively charged amino acid

THIS PAGE BLANK (USPTO)

within the hydrophobic portion of the LamB signal sequence. The mutant pre-LamB protein was not processed or secreted. In fact, cells producing the mutant precursor protein accumulated not only pre-LamB but also other unrelated precursor proteins in the cytoplasm. This finding suggested that the defective signal sequence still effectively competed with other signal sequences for interaction with component(s) of the secretory apparatus.

The Asp-substituted analog, like the other synthetic signal peptides, inhibited precursor protein processing in the mammalian system. In contrast, in the plant system, the Asp-substituted analog did not inhibit synthesis of either a secretory (pre-Prl) or cytoplasmic (globin) protein. This result suggests that insertion of a negative charge in the hydrophobic portion of a signal sequence interferes with the interaction of the signal region with the SRP-ribosome complex and thus fails to induce translation arrest.

The finding that a charge insertion in the signal sequence is disruptive to its activity in the wheat germ system, but is well tolerated in the reticulocyte lysate, lends further support for the suggestion that the signal region may interact with a component of the secretory apparatus other than or in addition to SRP in the mammalian system. Furthermore, the profile of biological properties displayed by the Asp-substituted signal sequence in the mammalian system *in vitro* corresponds to the properties observed for this kind of mutation *in vivo*: it does not inhibit translation, but still inhibits processing of its precursor and other nascent proteins (23). These observations further reinforce the concept that synthetic signal peptides can mimic the conformation of a native signal sequence as it emerges from the ribosome and are able to interact with components of the protein secretory apparatus.

The specificity of signal peptide studies was demonstrated by the use of a truncated signal peptide. This peptide failed to display any of the biological properties observed for the other signal peptides used. This finding is consistent with *in vivo* studies in which a minimum hydrophobic chain length (24) or an N-terminal positive charge (25, 26) on the signal sequence are required for protein translocation.

In conclusion, the diverse biological properties observed for the chemically synthesized signal peptide of pre-proparathyroid hormone and some of its analogs and the suggested presence of a factor(s) in addition to SRP involved in the secretory apparatus of the reticulocyte lysate system are consistent with the multiple and sequential functions thought to be served by signal regions *in vivo*. Signal regions are first thought to interact with SRP to promote a ribosome-SRP interaction, perhaps leading to translation arrest. However, once the ribosome-SRP complex interacts with a SRP receptor or docking protein on the RER, the signal region must dissociate from SRP in order to permit translocation of the nascent precursor protein across the RER membrane. This translocation process can occur in the absence of translation (27-29) and may be facilitated by the interaction of signal

peptide with other elements of the secretory apparatus. Given the complexity and sequential nature of these events, it is perhaps not unexpected that a subtly modified analog of the signal region might possess some, but not all, of the activities of an authentic signal region.

Acknowledgments—We wish to thank Jay Levy for the chemical synthesis of the synthetic peptides, Susan Fitzpatrick for amino acid analyses, Dr. Carl Bennett for peptide sequence analyses, John Smith for the mass spectroscopic analysis, and Dr. Robert Stein for critical reading of the manuscript.

REFERENCES

1. Walter, P., Gilmore, R., and Blobel, G. (1984) *Cell* 38, 5-8
2. Walter, P., Gilmore, R., and Blobel, G. (1985) in *Theories and Models in Cellular Transformation* (Santi, L., and Zardi, L., eds), pp. 113-120, Academic Press, London
3. Walter, P., Ibrahim, I., and Blobel, G. (1981) *J. Cell Biol.* 91, 545-550
4. Walter, P., and Blobel, G. (1981) *J. Cell Biol.* 91, 551-556
5. Walter, P., and Blobel, G. (1981) *J. Cell Biol.* 91, 557-561
6. Gilmore, R., Blobel, G., and Walter, P. (1982) *J. Cell Biol.* 95, 463-469
7. Gilmore, R., Walter, P., and Blobel, G. (1982) *J. Cell Biol.* 95, 470-477
8. Gilmore, R., and Blobel, G. (1983) *Cell* 35, 677-685
9. Meyer, D. I., Krause, E., and Dobberstein, B. (1982) *Nature* 297, 647-650
10. Meyer, D. I. (1985) *EMBO J.* 4, 2031-2033
11. Siegel, V., and Walter, P. (1985) *J. Cell Biol.* 100, 1913-1921
12. Majzoub, J. A., Rosenblatt, M., Fennick, B., Maunus, R., Kronenberg, H. M., Potts, J. T., Jr., and Habener, J. F. (1980) *J. Biol. Chem.* 255, 11478-11483
13. Austen, B. M., Hermon-Taylor, J., Kaderbhai, M. A., and Ridd, D. H. (1984) *Biochem. J.* 224, 317-325
14. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) *Biochemistry* 18, 5294-5299
15. Walter, P., and Blobel, G. (1983) *Methods Enzymol.* 96, 682-691
16. Walter, P., and Blobel, G. (1983) *Methods Enzymol.* 96, 84-93
17. Rosenblatt, M., Habener, J. F., Tyler, G. A., Shepard, G. L., and Potts, J. T., Jr. (1979) *J. Biol. Chem.* 254, 1414-1421
18. Laemmli, U. K. (1970) *Nature* 227, 680-685
19. Kurzhalia, T. V., Wiedmann, M., Girshovich, A. S., Bochkareva, E. S., Bielka, H., and Rapoport, T. A. (1986) *Nature* 320, 634-636
20. Krieg, U. C., Walter, P., and Johnson, A. E. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 8604-8608
21. Braunitzer, G., Hilse, K., Rudloff, V., and Hilschman, N. (1964) *Adv. Protein Chem.* 19, 1-71
22. Briggs, M. S., and Gierash, L. M. (1986) *Adv. Protein Chem.* 38, 110-180
23. Bankaitis, V. A., and Bassford, P. J., Jr. (1984) *J. Biol. Chem.* 259, 12193-12200
24. Bankaitis, V. A., Rasmussen, B. A., and Bassford, P. J., Jr. (1984) *Cell* 37, 243-252
25. Hall, M. N., Gabay, J., and Schwartz, M. (1983) *EMBO J.* 2, 15-19
26. Vlasuk, G. P., Inouye, S., Ito, H., Itakura, K., and Inouye, M. (1983) *J. Biol. Chem.* 258, 7141-7148
27. Mueckler, M., and Lodish, H. F. (1986) *Cell* 44, 629-637
28. Perara, E., Rothman, R. E., and Lingappa, V. R. (1986) *Science* 232, 348-352
29. Caulfield, M. P., Duong, L., and Rosenblatt, M. (1986) *J. Biol. Chem.* 261, 10953-10956

THIS PAGE BLANK (USPTO)

Volume 107
Number 6, Part 3
December 1988
Published by
The Rockefeller
University Press
and edited
in cooperation with
The American Society
for Cell Biology

The Journal of ^{FAL} Cell Biology

Take this
copy with
you to the
meeting.

Abstracts

The American Society for Cell Biology
and the American Society for
Biochemistry and Molecular Biology

Joint Meeting

29 January - 2 February 1989
San Francisco, California

THIS PAGE BLANK (USPTO)

4331 ISOLATION FROM BOVINE BRAIN OF A CALCIUM-BINDING PROTEIN THAT IS A CONSTITUENT OF A Ca-ATPase-CONTAINING ORGANELLE. R. J. Johnson, Benson, and R.E. Fine^{1,2}, ¹Boston University School of Medicine, ²Edith Morse Rogers Memorial VA Hospital, G.R.E.C.C.

Using a Ca-oxalate density shift method we have isolated a membrane fraction from bovine brain that contains a CaATPase and a 58,000 M_r protein (CBP-58) similar in some respects to muscle calsequestrin (Csq). We have isolated CBP-58 using anion exchange and hydrophobic chromatography. Like Csq, it has an acidic pI (4.7) containing 10% glu and asp. By SDS-PAGE at pH 8.8 its M_r is 58 kD, while at pH 7, it is approximately 52 kD. When SDS-PAGE are stained with the carbocyanine dye, Stains-All, CBP-58 appears as a blue band. This color is not due to the presence of sialic acid residues or phosphate groups. CBP-58 elutes from DEAE-cellulose between 350 and 425 mM NaCl. Unlike Csq, CBP-58 binds to phenyl-sepharose in the presence of 10 mM CaCl₂, and elutes with EGTA. The amino acid composition of this protein is similar to neither Csq, CAB-61 (calsequestrin), nor the calalectrins.

In terms of its cell biology, when brain microsomes are incubated in the presence of potassium oxalate, CaCl₂, MgCl₂, and ATP, CBP-58 sediments with membranes that contain a CaATPase. It is released from these membranes by neither freezing and thawing, nor by washing in 1 M NaCl; but by treatment with 0.1 M Na₂CO₃ at pH 11.3. When brains microsomes are treated with trypsin, a blue-staining band of 58 kD is spared digestion unless Triton X-100 is present. We propose that CBP-58 is a constituent of the lumen of the smooth endoplasmic reticulum, and that it serves to bind calcium within the organelle which in axons is known to sequester calcium. This work is supported by NIA AG05894-16 to REF.

4333 Is there a BiP-like protein in the endoplasmic reticulum of plant cells?

J.A. Mierzyk, Seed Biosynthesis Research Unit, USDA, ARS, Northern Regional Research Center, 1815 North University St., Peoria, IL 61604

Immunoglobulin heavy-chain binding protein (BiP, also called GRP-78) is a member of the HSP-70 family of heat-shock related proteins. Within the lumen of the ER there is a tight binding of BiP to incompletely or incorrectly assembled proteins. Due in part to its abundance in secretory cells, it has also been suggested that BiP may play a role in the normal assembly of secreted proteins. Maize endosperm cultures are quite active in the extracellular secretion of acid hydrolases, and electron microscopy of the cultured cells revealed a well developed endomembrane system. This combination of features provides an excellent system for molecular characterization of the secretory pathway in plant cells. One aspect of the characterization is the search for a BiP-like protein. Maize endosperm microsomes analyzed by SDS-PAGE, contain an abundant soluble protein of Mr 76,000; slightly smaller than mammalian BiP. The maize protein will bind to immobilized ATP, a characteristic of BiP and other HSP-related proteins. Monoclonal antibodies against mouse BiP were used to probe total microsomal proteins from rat liver and maize endosperm. These antibodies gave a positive signal with BiP from the rat membranes but not with any protein from maize. The effects of treatment with A23187 or tunicamycin, compounds that stimulate the synthesis of BiP in mammalian systems, will be reported.

4335 Coupled Induction of Secretory Proteins and Intracellular Compartments Involved in the Secretory Pathway in AR4-2J Cells by Glucocorticoids. G.A. Scheele, U. Steinhilber, B. Swarovsky, and H.F. Kern, Lab. of Cell and Molecular Biology, The Rockefeller University, New York, NY 10021 and Dept. of Cell Biology, Philipps University, Marburg, FRG (Spon. by G. Scheele).

Treatment of AR4-2J cells with 10 nM dexamethasone for 96 hrs led to an increase in intracellular compartments responsible for synthesis, packaging and storage of secretory proteins and to the formation of gap junctions. The 30% increase in cell size was exclusively due to cytoplasmic enlargement. Although the projection area of the Golgi complex did not significantly change, a 30-fold increase in the area of rough endoplasmic reticulum was observed after dexamethasone treatment. Secretion granules, shown to contain pancreatic enzymes by immunocytochemistry, were observed only after treatment with dexamethasone. The observed 2-fold increase in total protein synthesis was accounted for by increases in the synthesis of pancreas-specific enzymes: trypsinogens, 15-fold; amylase, 10-fold; procarboxypeptidases, 8-fold; and chymotrypsinogens, 4-fold. A 4-5 fold increase in mRNA concentrations was observed by dot-blot hybridization of total cellular RNA using cDNAs for amylase, anionic trypsinogen and lipase as probes. The data indicate that dexamethasone in AR4-2J cells may act on both the transcriptional and translational levels. Recent data on the early effects of hormone treatment show an increase of mRNA levels for the secretory enzymes within 3-6 hrs after exposure to dexamethasone. We conclude that the pancreatic acinar AR4-2J cell line is a suitable model for studying the mechanisms involved in differentiation of pancreatic epithelial cells.

4332 Staining of the endoplasmic reticulum by the fluorescent dye DiOC6(3) and by an antibody to BiP. M. Teraski and T.S. Reese, Laboratory of Neurobiology, NINCDS, NIH at Woods Hole.

Individual cells of a monkey kidney epithelial cell line (CV-1) were fixed lightly with 0.025% glutaraldehyde, stained by the fluorescent dye DiOC6(3), photographed, post-fixed with cold methanol then processed for immunofluorescence with a rat monoclonal antibody to BiP (Bole et al., J. Cell Biol. 102:1558), a protein in the ER. The peripheral networks co-localized by the dye and by immunofluorescence were identical. Since BiP is present in the rough but not smooth ER of lymphocytes and liver cells (Bole et al. in preparation), these results suggest that the ER in the periphery of CV-1 cells is rough ER. Thus, the ER in the CV-1 and, possibly, other cells is pervasively distributed. Since essentially all membranes (both lipid and protein components) are thought to be synthesized in the ER, the ER can be thought of as a membrane synthesizing compartment distinct from the endo- and exocytic pathways. If this pre-Golgi membrane synthesizing compartment forms as and remains a continuous compartment, then it will grow by expansion, accounting for the continuity of ER membranes we have defined in CV-1 cells. Extension of these membranes may depend on attached motors which move towards the + end of microtubules (Teraski et al., J. Cell Biol. 103:1557, Dabora and Sheetz, Cell 54:27). The resulting single, interconnected compartment distributed throughout the cell could play a role in regulating and coordinating activities at the level of the whole cell by, for example, its ability to transport ions or other molecules across its membranes which could then diffuse to other regions of the cell.

4334 Are There Two Distinct Protein Disulfide Isomerases in the ER? M. Srinivasan, S. Haugejorden, R.A. Mazzarella, and M. Green, Saint Louis University School of Medicine, St. Louis, MO 63104

We have isolated full-length, expressible cDNA clones for murine protein disulfide isomerase (PDI) and Erp72, an abundant luminal protein of the ER. A comparison of the sequences of these two proteins revealed that Erp72 contains three copies of the sequences that are thought to function as the active sites of PDI. Thus, Erp72 appears to be a newly identified member of the PDI family. At present, we are trying to determine whether Erp72 has protein disulfide isomerase activity. PDI possesses the C-terminal sequence KDEL, while the C-terminal sequence of Erp72 is KEEL. In order to determine if these sequences act as ER retention sequences for these proteins, we have constructed mutants of PDI and Erp72 which lack these sequences. The sorting fate of the mutant proteins is being determined in transient expression studies in COS cells.

4336 The Structure and Sorting of Erp99, an Abundant Glycoprotein of the RER. R.A. Mazzarella and M. Green, Saint Louis University School of Medicine, St. Louis, MO 63104

Erp99 is homologous to the 90kDa heat shock protein (hsp90) of *S. cerevisiae* and to the 94kDa glucose regulated protein (GRP94) of mammalian cells. Our model for the orientation of Erp99 in the ER membrane places the NH₂-terminus in the lumen of the ER, predicts one membrane spanning, stop transfer region and places most of Erp99 on the cytoplasmic side of the ER membrane. Erp99 possesses the C-terminal sequence KDEL, the ER retention signal for heavy chain binding protein. In order to test our model and assess the role of KDEL, we have constructed several deletion mutants of Erp99 and determined how these mutants are sorted in COS cells. The results of these experiments are entirely consistent with our model. The proposed luminal domain fragmentation can be secreted if the proposed transmembrane region is removed. The putative cytoplasmic domain fragment is capable of entering the lumen and being glycosylated if the signal sequence is attached to it. Furthermore, the removal of KDEL or the entire hydrophilic C-terminal sequence containing KDEL, does not result in the secretion of Erp99. Thus, KDEL is either not an ER retention signal for Erp99 or is not the only retention signal in this protein.

THIS PAGE BLANK (USPTO)

4331 ISOLATION FROM BOVINE BRAIN OF A CALCIUM-BINDING PROTEIN THAT IS A CONSTITUENT OF A Ca-ATPase-CONTAINING ORGANELLE. R. J. Johnson, Banson¹ and R. E. Fine². ¹Boston University School of Medicine, ²Edith Morse Rogers Memorial VA Hospital, G.R.E.C.C.

Using a Ca-oxalate density shift method we have isolated a membrane fraction from bovine brain that contains a CaATPase and a 58,000 M_r protein (CBP-58) similar in some respects to muscle calsequestrin (Csq). We have isolated CBP-58 using anion exchange and hydrophobic chromatography. Like Csq, it has an acidic pI (4.7) containing 30% glu and asp. By SDS-PAGE at pH 8.8 its M_r is 58 kD, while at pH 7, it is approximately 52 kD. When SDS-PAGE are stained with the carbocyanine dye, Stains-All, CBP-58 appears as a blue band. This color is not due to the presence of sialic acid residues or phosphate groups. CBP-58 elutes from DEAE-cellulose between 350 and 425 mM NaCl. Unlike Csq, CBP-58 binds to phenyl-sepharose in the presence of 10 mM $CaCl_2$ and elutes with EGTA. The amino acid composition of this protein is similar to neither Csq, CAB-63 (calregulin), nor the calelectrins.

In terms of its cell biology, when brain microsomes are incubated in the presence of potassium oxalate, $CaCl_2$, $MgCl_2$ and ATP, CBP-58 sediments with membranes that contain a CaATPase. It is released from these membranes by neither freezing and thawing, nor by washing in 1 M NaCl; but by treatment with 0.1 M Na_2CO_3 at pH 11.3. When brains microsomes are treated with trypsin, a blue-staining band of 58 kD is spared digestion unless Triton X-100 is present. We propose that CBP-58 is a constituent of the lumen of the smooth endoplasmic reticulum, and that it serves to bind calcium within the organelle which in axons is known to sequester calcium. This work is supported by NIA AG05894-16 to REF.

4333 Is there a BiP-like protein in the endoplasmic reticulum of plant cells?

J.A. Miernyk, Seed Biosynthesis Research Unit, USDA, ARS, Northern Regional Research Center, 1815 North University St., Peoria, IL 61604

Immunoglobulin heavy-chain binding protein (BiP, also called GRP-78) is a member of the HSP-70 family of heat-shock related proteins. Within the lumen of the ER there is a tight binding of BiP to incompletely or incorrectly assembled proteins. Due in part to its abundance in secretory cells, it has also been suggested that BiP may play a role in the normal assembly of secreted proteins. Maize endosperm cultures are quite active in the extracellular secretion of acid hydrolases, and electron microscopy of the cultured cells revealed a well developed endomembrane system. This combination of features provides an excellent system for molecular characterization of the secretory pathway in plant cells. One aspect of the characterization is the search for a BiP-like protein. Maize endosperm microsomes analyzed by SDS-PAGE, contain an abundant soluble protein of M_r 76,000; slightly smaller than mammalian BiP. The maize protein will bind to immobilized ATP, a characteristic of BiP and other HSP-related proteins. Monoclonal antibodies against mouse BiP were used to probe total microsomal proteins from rat liver and maize endosperm. These antibodies gave a positive signal with BiP from the rat membranes but not with any protein from maize. The effects of treatment with A23187 or tunicamycin, compounds that stimulate the synthesis of BiP in mammalian systems, will be reported.

4332 Staining of the DiOC6(3) and by an antibody
Laboratory of Neurobiology, N

Individual cells of a m fixed lightly with 0.025% gl DiOC6(3), photographed, pos for immunofluorescence with al., J. Cell Biol. 102:1558). co-localized by the dye a Since BiP is present in the liver cells (Bole et al, in bodies but not axons (Bridge the ER in the periphery of CV-1 and, possibly, othe essentially all membranes thought to be synthesized i membrane synthesizing co exocytic pathways. If compartment forms as and will grow by expansion, acc we have defined in CV-1 depend on attached mo microtubules (Terasaki et al., Cell 54:27). The rest distributed throughout the coordinating activities at the ability to transport ions or could then diffuse to other

4334 Are There Two C ER? M. Srinivasan, S. Ha
Saint Louis University Sch

We have isolated full-l protein disulfide isomerase protein of the ER. A comp proteins revealed that ERp sequences that are thought Thus, ERp72 appears to b family. At present, we are protein disulfide isomerase sequence KDEL, while the In order to determine if th sequences for these prote and ERp72 which lack the mutant proteins is being c in COS cells.

THIS PAGE BLANK (USPTO)

nature

INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

Volume 342 No. 6245 2 November 1993 \$4.95



ANTIBODY PRODUCTION IN TRANSGENIC PLANTS

CELL BIOLOGY
product review

EXHIBIT D

THIS PAGE BLANK (USPTO)

SAN DIEGO COUNTY

Los Angeles Times

CIRCULATION
LITHO DAILY / 6,017,791 SUNDAY

THURSDAY, NOVEMBER 2, 1989

CEN. C. 1989 / THE TIMES PUBLISHING COMPANY / 100,000 PAGES

DAILY 25¢

DESIGNATED A READING ROOM

COLUMN ONE

Parents
Spell the
Difference

■ Their extraordinary role helped bring an end to Beverly Hills teachers strike. Educators call it a how-to manual.

By JEAN MERI
TIMES EDUCATION WRITER

When Day 3 of the Beverly Hills teachers strike dawned with no negotiations in sight, real estate investor Robert Schwab, whose children attend the city's prestigious public schools, persuaded the union and the district to go back to the bargaining table.

Two other fathers, Eli Blumenfeld and Albert Gerson, met with representatives of both sides and determined that money was the missing ingredient, so they worked with other parents to raise hundreds of thousands of dollars in cash and pledges to put toward teacher salaries.

Another parent, real estate attorney Tom Levay, building on the groundwork that Schwab laid, put together a negotiating session last weekend that would lead to a settlement in the district's first-ever teacher strike.

Throughout a city known for the excellence of its public schools, parents met, offered ideas and professional skills, took on their checkbooks and prodded their elected officials.

On Wednesday, members of the local teachers union approved a settlement giving them a 12% salary increase over two years and ending the walkout that began Oct. 16. As teachers prepare to go back to their classrooms today, the parents of Beverly Hills' students are

Educators around the country



Andrew Hatt of Scripps Clinic in La Jolla checks a petri dish for root growth in antibody-producing plant.

Scripps Technique Seen
as Big Medical Stride

■ Science: Researchers develop method of producing antibodies that could open up new medical field.

By THOMAS H. MAUGH II
TIMES SCIENCE WRITER

Researchers at the Scripps Clinic in La Jolla have developed an inexpensive technique for producing specialized antibodies that promises to open up a whole new area of medicine, including cancer therapy, and to greatly expand medical diagnostics.

The technique also could be used to give agricultural crops a functioning "immune system" that would protect them against insects, fungi and other pathogens and to create new plants that could be

used to clean up polluted water.

The use of antibodies was revolutionized in 1975 by the discovery of techniques for producing large quantities of a single antibody. Georges Kohler and Cesar Milstein of the Laboratory for Molecular Biology in Cambridge, England, isolated single, short-lived white blood cells that produced one antibody and fused them with long-lived cancer cells.

The resulting cells, called hybridomas, inherit immortality—but

the capacity of what they had been

Conferees Said
to Retain B-2,
Cut 'Star Wars'

By JOHN M. BRODER
TIMES STAFF WRITER

WASHINGTON—House and Senate conferees Wednesday reached accord on virtually all remaining issues involving the \$305-billion 1990 defense budget, including the "Star Wars" antimissile program but keeping the costly and controversial B-2 Stealth bomber alive, congressional sources said.

Although a number of relatively small questions remained unresolved Wednesday night, committee leaders are likely to announce the budget agreement today.

Although a number of relatively

Truce Is Over,
Ortega Declares

■ Nicaragua: He vows to drive out the Contras. He also calls for peace talks, but their resumption is uncertain.

By RICHARD BOLDREAU
TIMES STAFF WRITER

MANAGUA, Nicaragua—President Daniel Ortega formally ended a 19-month-old cease-fire against U.S.-backed rebels Wednesday and said the Sandinista army will try to push them back to bases in Honduras. But he asked rebel leaders to meet in New York next Monday for talks on ending the newly resurgent Nicaraguan war.

Enrique Bermudez, the senior Contra commander, said he will accept the offer if Nicaragua's Roman Catholic cardinal is also invited, a condition that Ortega quickly accepted.

The collapse of the cease-fire plunged this war-weary country into its deepest period of uncertainty in two years of peace efforts. Cardinal Miguel Obando y Bravo appealed for "restraint and maturity" by both armies, warning that intensified fighting could disrupt the elections scheduled next Feb. 28.

A resumption of peace talks, broken off in June, 1988, is far from certain, Ortega insisted that Honduras be represented, but that government balked at agreeing to an U.N. Secretary General Javier Perez de Cuellar, the would-be host, was described by aides as surprised by Ortega's announcement and would not confirm that a meeting will take place.

"This is extremely unusual," a U.N. official said in New York. "It might have been easier if the Sandinistas had notified us first."

Ortega's army, by the way, the



Ortega at press conference

U.S. Condemns
Move, Hopes It
Will Backfire

By DAVID LAUTER
TIMES STAFF WRITER

WASHINGTON—U.S. officials denounced Nicaraguan President Daniel Ortega on Wednesday for ending the 19-month cease-fire with the Contras, but they expressed optimism that the move will backfire and inflame anti-Sandinista sentiments.

White House Press Secretary Martin Pitzwater said Ortega's decision is "deplorable," adding that it confirms Ortega's "desire to maintain power at all costs."

Last Friday, Administration officials expressed shock and anger at Ortega's threat to resume the

THIS PAGE BLANK (USPTO)

SHORT COMMUNICATION

Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites

Henrik Nielsen, Jacob Engelbrecht¹, Søren Brunak and Gunnar von Heijne²

Center for Biological Sequence Analysis, Department of Chemistry, The Technical University of Denmark, DK-2800 Lyngby, Denmark and ²Department of Biochemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden

¹Present address: Novo Nordisk A/S, Scientific Computing, Building 9M1, Novo Alle, DK-2880 Bagsværd, Denmark

We have developed a new method for the identification of signal peptides and their cleavage sites based on neural networks trained on separate sets of prokaryotic and eukaryotic sequence. The method performs significantly better than previous prediction schemes and can easily be applied on genome-wide data sets. Discrimination between cleaved signal peptides and uncleaved N-terminal signal-anchor sequences is also possible, though with lower precision. Predictions can be made on a publicly available WWW server.

Keywords: cleavage sites/protein sorting/secretion/signal peptide

Introduction

Signal peptides control the entry of virtually all proteins to the secretory pathway, both in eukaryotes and prokaryotes (Gierasch, 1989; von Heijne, 1990; Rapoport, 1992). They comprise the N-terminal part of the amino acid chain and are cleaved off while the protein is translocated through the membrane. The common structure of signal peptides from various proteins is commonly described as a positively charged n-region, followed by a hydrophobic h-region and a neutral but polar c-region. The (-3,-1) rule states that the residues at positions -3 and -1 (relative to the cleavage site) must be small and neutral for cleavage to occur correctly (von Heijne, 1983, 1985).

A strong interest in the automated identification of signal peptides and the prediction of their cleavage sites has been evoked not only by the huge amount of unprocessed data available, but also by the industrial need to find more effective vehicles for the production of proteins in recombinant systems. The most widely used method for predicting the location of the cleavage site is a weight matrix which was published in 1986 (von Heijne, 1986). This method is also useful for discriminating between signal peptides and non-signal peptides by using the maximum cleavage site score. The original matrices are commonly used today, even though the amount of signal peptide data available has increased since 1986 by a factor of 5-10.

Here, we present a combined neural network approach to the recognition of signal peptides and their cleavage sites, using one network to recognize the cleavage site and another network to distinguish between signal peptides and non-signal peptides. A similar combination of two pairs of networks has been used with success to predict the intron splice sites

in pre-mRNA from humans and the dicotyledonous plant *Arabidopsis thaliana* (Brunak *et al.*, 1991; S.Hebsgaard, P.Korning, J.Engelbrecht, P.Rouze and S.Brunak, submitted). Artificial neural networks have been used for many biological sequence analysis problems (Hirst and Sternberg, 1992; Presnell and Cohen, 1993). They have also been applied to the twin problems of predicting signal peptides and their cleavage sites, but until now without leading to practically applicable prediction methods with significant improvements in performance compared with the weight matrix method (Arrigo *et al.*, 1991; Ladunga *et al.*, 1991; Schneider and Wrede, 1993).

Materials and methods

The data were taken from SWISS-PROT version 29 (Bairoch and Boeckmann, 1994). The data sets were divided into prokaryotic and eukaryotic entries and the prokaryotic data sets were further divided into Gram-positive eubacteria (*Firmicutes*) and Gram-negative eubacteria (*Gracilicutes*), excluding *Mycoplasma* and *Archaeobacteria*. Viral, phage and organellar proteins were not included. In addition, two single-species data sets were selected, a human subset of the eukaryotic data and an *Escherichia coli* subset of the Gram-negative data.

The sequence of the signal peptide and the first 30 amino acids of the mature protein from the secretory protein were included in the data set. The first 70 amino acids of each sequence were used from the cytoplasmic and (for the eukaryotes) nuclear proteins. In addition, a set of eukaryotic signal anchor sequences, i.e. N-terminal parts of type II membrane proteins (von Heijne, 1988), were extracted (see Figure 1).

As an example of a large-scale application of the finished method, we used the *Haemophilus influenzae* Rd genome—the first genome of a free-living organism to be completed (Fleischmann *et al.*, 1995). We have downloaded the sequences of all the predicted coding regions in the *H.influenzae* genome from the World Wide Web (WWW) server of the Institute for Genomic Research at <http://www.tigr.org/>. Only the first 60 positions of each sequence were analysed.

We have attempted to avoid signal peptides where the cleavage sites are not experimentally determined, but we are not able to eliminate them completely, since many database entries simply lack information about the quality of the evidence. The details of the data selection are described in the WWW server and in an earlier paper (Nielsen *et al.*, 1996a).

Redundancy in the data sets was avoided by excluding pairs of sequences which were functionally homologous, i.e. those that had more than 17 (eukaryotes) or 21 (prokaryotes) exact matches in a local alignment (Nielsen *et al.*, 1996a). Redundant sequences were removed using an algorithm which guarantees that no pairs of homologous sequences remain in the data set (Hobohm *et al.*, 1992). This procedure removed 13-56% of the sequences. The numbers of non-homologous sequences remaining in the data sets are shown in Table I. Redundancy

THIS PAGE BLANK (USPTO)

Table I. Data and performance values

Source	Data		Network architecture (window/hidden units)		Performance	
	(Number of sequences)		C-score	S-score	Cleavage site location (% correct)	Signal peptide discrimination (correlation)
	Signal peptides	Non-secretory proteins				
Human	416	251	15+4/2	27 / 4	68.0 (67.9)	0.96 (0.97)
Eukaryote	1011	820	17+2/2	27 / 4	70.2	0.97
<i>E.coli</i>	105	119	15+2/2	39 / 0	83.7 (85.7)	0.89 (0.92)
Gram-	266	186	11+2/2	19 / 3	79.3	0.88
Gram+	141	64	21+2/0	19 / 3	67.9	0.96

Data: the number of sequences of signal peptides and non-secretory (i.e. cytoplasmic or nuclear) proteins in the data sets after redundancy reduction. The organism groups are eukaryotes, human, Gram-negative bacteria ('Gram-'), *E.coli* and Gram-positive bacteria ('Gram+'). The human data are subsets of the eukaryotic data and the *E.coli* data are subsets of the Gram-negative data. The signal anchor and *H.influenzae* data are not shown in the table. *Network architecture*: the size of the input window and the number of hidden computational units ('neurons') in the optimal neural networks chosen for each data set. *C-score* networks have asymmetrical input windows. *Performance*: the percentage of signal peptide sequences where the cleavage site was predicted to be at the correct location according to the maximal value of the Y-score (see Figure 2). The ability of the method to distinguish between the signal peptides and the N-terminals of non-secretory proteins (based on the mean value of the S-score in the region between position 1 and the predicted cleavage site position) is measured by the correlation coefficients (Mathews, 1975). Both performance values are measured on the test sets (the average of five cross-validation tests). The values given in parentheses indicate the performance for the human sequences when using networks trained on all eukaryotic data and for the *E.coli* sequences when using Gram-negative networks respectively.

reduction was not applied to the signal anchor data or the *H.influenzae* data, since these were not used as training data.

Neural network algorithms

The signal peptide problem was posed to the neural networks in two ways: (i) recognition of the cleavage sites against the background of all other sequence positions and (ii) classification of amino acids as belonging to the signal peptide or not. In the latter case, negative examples included both the first 70 positions of non-secretory proteins and the first 30 positions of the mature part of secretory proteins.

The neural networks were feed-forward networks with zero or one layer of two to 10 hidden units, trained using back-propagation (Rumelhart *et al.*, 1986) with a slightly modified error function. The sequence data were presented to the network using sparsely encoded moving windows (Qian and Sejnowski, 1988; Brunak *et al.*, 1991). Symmetric and asymmetric windows of a size varying from five to 39 positions were tested.

Based on the numbers of correctly and incorrectly predicted positive and negative examples, we calculated the correlation coefficient (Mathews, 1975). The correlation coefficients of both the training and test sets were monitored during training and the performance of the training cycle with the maximal test set correlation was recorded for each training run. The networks chosen for inclusion in the WWW server have been trained until this cycle only.

The test performances have been calculated by cross-validation: each data set was divided into five approximately equal-sized parts and then every network run was carried out with one part as test data and the other four parts as training data. The performance measures were then calculated as an average over the five different data set divisions.

For each of the five data sets, one signal peptide/non-signal peptide network architecture and one cleavage site/non-cleavage site network architecture was chosen on the basis of the test set correlation coefficients. We did not pick the architecture with absolutely the best performance, but instead the smallest network that could not be significantly improved by enlarging the input window or adding more hidden units.

The trained networks provide two different scores between zero and one for each position in an amino acid sequence. The output from the signal peptide/non-signal peptide networks, the S-score, can be interpreted as an estimate of the probability of the position belonging to the signal peptide, while the output from the cleavage site/non-cleavage site networks, the C-score, can be interpreted as an estimate of the probability of the position being the first in the mature protein (position + 1 relative to the cleavage site).

If there are several C-score peaks of comparable strength, the true cleavage site may often be found by inspecting the S-score curve in order to see which of the C-score peaks coincides best with the transition from the signal peptide to the non-signal peptide region. In order to formalize this and improve the prediction, we have tried a number of linear and non-linear combinations of the raw network scores and evaluated the percentage of sequences with correctly placed cleavage sites in the five test sets. The best measure was the geometric average of the C-score and a smoothed derivative of the S-score, termed the Y-score:

$$Y_i = \sqrt{C_i \Delta_d S_i} \quad (1)$$

where $\Delta_d S_i$ is the difference between the average S-score of d positions before and d position after position i :

$$\Delta_d S_i = \frac{1}{d} \left(\sum_{j=1}^d S_{i-j} - \sum_{j=0}^{d-1} S_{i+j} \right) \quad (2)$$

In Figure 2(A), examples of the values of the C-, S- and Y-scores are shown for a typical signal peptide with a typical cleavage site. The C-score has one sharp peak that corresponds to an abrupt change in the S-score from a high to low value. Among the real examples, the C-score may exhibit several peaks and the S-score may fluctuate. We define a cleavage site as being correctly located if the true cleavage site position corresponds to the maximal Y-score (combined score).

For a typical non-secretory position, the values of the C-, S- and Y-scores are lower, as shown in Figure 2(B). We found the best discriminator between signal peptides and non-secretory

THIS PAGE BLANK (USPTO)

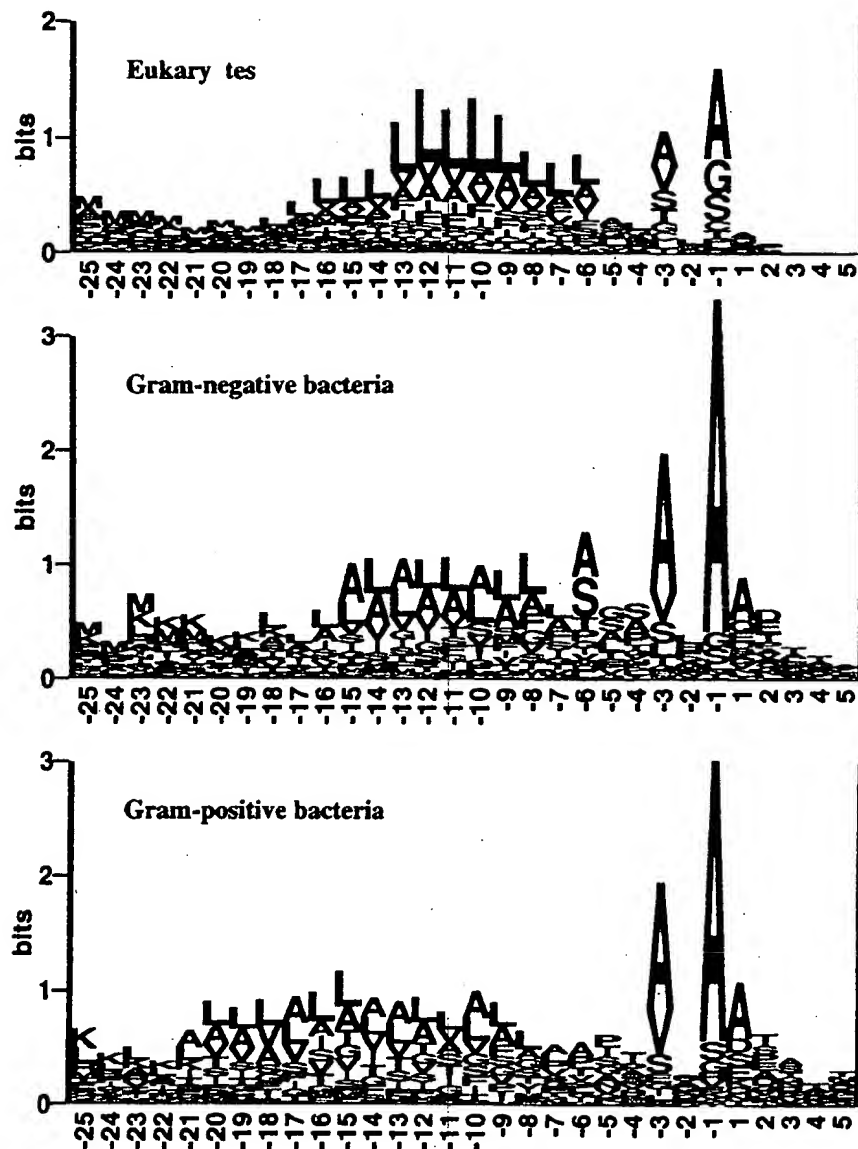


Fig. 1. Sequence logos (Schneider and Stephens, 1990) of signal peptides, aligned by their cleavage sites. The total height of the stack of letters at each position shows the amount of information, while the relative height of each letter shows the relative abundance of the corresponding amino acid. The information is defined as the difference between the maximal and actual entropy (Shannon, 1948): $I_j = H_{\max} - H_j = \log_2 20 + \sum_{\alpha} n_j(\alpha)/N_j \log_2 n_j(\alpha)/N_j$, where $n_j(\alpha)$ is the number of occurrences of the amino acid α and N_j is the total number of letters (occupied positions) at position j . Positively and negatively charged residues are shown in blue and red respectively, while uncharged polar residues are green and hydrophobic residues are black.

proteins to be the average of the S-score in the predicted signal peptide region, i.e. from position 1 to the position immediately before the position where the Y-score has a maximal value. If this value—the mean S-score—is greater than 0.5, we predict the sequence in question to be a signal peptide (cf. Figure 3).

The relationship between the various performance measures and their development during the training process is described in detail elsewhere (Nielsen *et al.*, 1997).

Results and discussion

The optimal network architecture and corresponding predictive performance for all the data sets are shown in Table I. The C-

score problem is best solved by networks with asymmetric windows, i.e. windows including more positions upstream than downstream of the cleavage site. This corresponds well with the location of the cleavage site pattern information which is shown as sequence logos (Schneider and Stephens, 1990) in Figure 1. The S-score problem, on the other hand, is best solved by symmetric or approximately symmetric windows.

Although our method is able to locate cleavage sites and discriminate signal peptides from non-secretory proteins with a reasonably high reliability, the accuracy of the cleavage site location is lower than that reported for the original weight matrix method (von Heijne, 1986): 78% for eukaryotes and

THIS PAGE BLANK (USPTO)

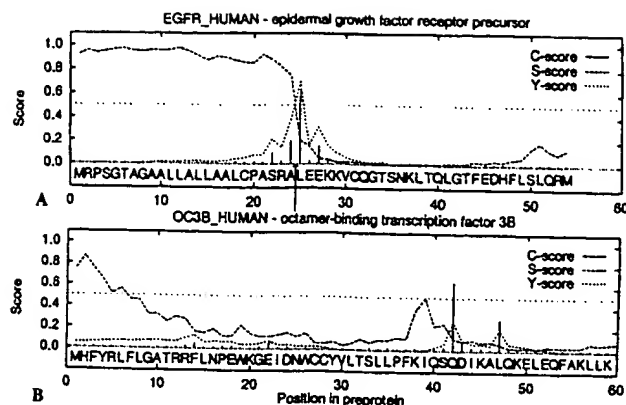


Fig. 2. Examples of network output. The values of the C- (output from cleavage site networks), S- (output from signal peptide networks) and Y-scores (combined cleavage site score, $Y_k = \sqrt{C_k \Delta_k S_k}$) are shown for each position in the sequence. The C- and S-scores are averages over five networks trained on different parts of the data. Note: the C- and Y-scores are high for the position immediately after the cleavage site, i.e. the first position in the mature protein. (A) A successfully predicted signal peptide. The true cleavage site is marked with an arrow. (B) A non-secretory protein. For many non-secretory proteins, all three scores are very low throughout the sequence. In this example, there are peaks of the C- and S-scores, but the sequence is still easily classified as non-secretory, since the C-score peak occurs far away from the S-score decline and the region of the high S-score is far too short.

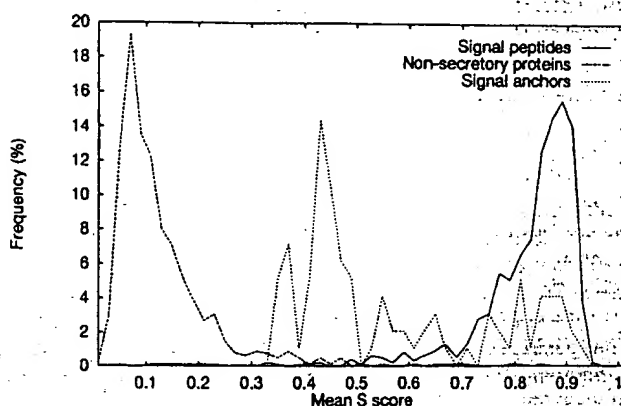


Fig. 3. Distribution of the mean signal peptide score (S-score) for signal peptides and non-signal peptides (eukaryotic data only). 'Non-secretory proteins' refer to the N-terminal parts of cytoplasmic or nuclear proteins, while 'signal anchors' are the N-terminal parts of type II membrane proteins. The mean S-score of a sequence is the average of the S-score over all positions in the predicted signal peptide region (i.e. from the N-terminal to the position immediately before the maximum of the Y-score). The bin size of the distribution is 0.02.

89% for prokaryotes (not divided into Gram-positive and -negative). When the original weight matrix is applied to our recent data set, however, the performance is much lower. This suggests a larger variation in the examples of the signal peptides found since then. It may, of course, also reflect a higher occurrence of errors in our automatically selected data than in the manually selected 1986 set.

In order to compare the strength of the neural network approach to the weight matrix method, we recalculated new weight matrices from our new data and tested the performances of these (results not shown). The weight matrix method was comparable to the neural networks when calculating the C-score, but was practically unable to solve the S-score problem

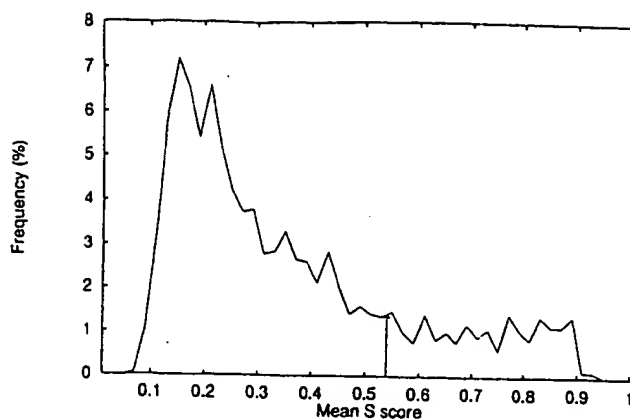


Fig. 4. Distribution of the mean signal peptide score (S-score) for all the predicted *H. influenzae* coding sequences. The mean S-score is calculated using networks trained on the Gram-negative data set. The bin size of the distribution is 0.02. The arrow shows the optimal cut-off for predicting a cleavable signal peptide. The predicted number of secretory proteins in *H. influenzae* (corresponding to the area under the curve to the right of the arrow) is 330 out of 1680 (20%).

and therefore did not provide the possibility of calculating the combined Y-score.

Note that the prediction performances reported here correspond to minimal values. The test sets in the cross-validation have a very low sequence similarity; in fact, the sequence similarity is so low that the correct cleavage sites cannot be found by alignment (Nielsen *et al.*, 1996a). This means that the prediction accuracy on sequences with some similarity to the sequences in the data sets will in general be higher.

The differences between the signal peptides from different organisms are apparent from Figure 1. The signal peptides from Gram-positive bacteria are considerably longer than those of other organisms, with much more extended h-regions, as observed previously (von Heijne and Abrahmsén, 1989). The prokaryotic h-regions are dominated by Leu (L) and Ala (A) in approximately equal proportions and in the eukaryotes they are dominated by Leu with some occurrence of Val (V), Ala, Phe (F) and Ile (I). Close to the cleavage site, the (-3,-1) rule is clearly visible for all three data sets, but while a number of different amino acids are accepted in the eukaryotes, the prokaryotes accept alanine almost exclusively in these two positions. In the first few positions of the mature protein (downstream of the cleavage site) the prokaryotes show certain preferences for Ala, negatively charged (D or E) amino acids, and hydroxy amino acids (S or T), while no pattern can be seen for the eukaryotes. In the leftmost part of the alignment, the positively charged residue Lys (K) [and to a smaller extent Arg (R)] is seen in the prokaryotes, while the eukaryotes show a somewhat weaker occurrence of Arg (barely visible in the figure) and almost no Lys. This corresponds well with the hypothesis that positive residues are required in the n-region where the N-terminal Met is formulated for prokaryotes, but not necessarily for eukaryotes where the N-terminal Met in itself carries a positive charge (von Heijne, 1985).

The difference in structure is reflected in the performances of the trained neural networks (see Table I). Gram-negative cleavage sites have the strongest pattern—i.e. the highest information content—and, consequently, they are the easiest to predict, both at the single-position and at the sequence level. The eukaryotic cleavage sites are significantly more difficult

THIS PAGE BLANK (USPTO)

to predict. Gram-positive cleavage sites are slightly more difficult to predict than the eukaryotic ones, which would not be expected from the sequence logos (Figure 1), since they show nearly as high an information content as the Gram-negative cleavage sites, but the longer Gram-positive signal peptides means that the cleavage sites have to be located against a larger background of non-cleavage site positions. The discrimination of signal peptides versus non-secretory proteins, on the other hand, is better for the eukaryotes than for the prokaryotes. This may be due to the more characteristic leucine-rich h-regions of the eukaryotic signal peptides.

The logos for the human and *E.coli* data sets are not shown, since they show no significant differences from those of the eukaryotes or Gram-negative bacteria respectively. Accordingly, the predictive performance was not improved by training the networks on single-species data sets. On the contrary, the *E.coli* signal peptides are predicted even better by the Gram-negative networks than by the *E.coli* networks (probably due to the relatively small size of the *E.coli* data set). In other words, we have found no evidence for species-specific features of the signal peptides of humans and *E.coli*.

Signal anchors often have sites similar to signal peptide cleavage sites after their hydrophobic (transmembrane) region. Therefore, a prediction method can easily be expected to mistake signal anchors for peptides. In Figure 3, the distribution of the mean S-score for the 97 eukaryotic signal anchors is included. It shows some overlap with the signal peptide distribution. If the standard cut-off of 0.5 is applied to the signal anchor data sets, 50% of the eukaryotic signal anchor sequences are falsely predicted as signal peptides (the corresponding figure for the human signal anchors is 75% when using human networks and 68% when using eukaryotic networks). With a cut-off optimized for signal anchor versus signal peptide discrimination (0.62), we were able to lower this error rate to 45% for the eukaryotic data set. The mean S-score still gives a better separation than the maximal C- or Y-score, which indicates that the pseudo-cleavage sites are in fact rather strong.

However, the pseudo-cleavage sites often occur further from the N-terminal than genuine cleavage sites do. If we do not accept signal peptides longer than 35 residues (this will exclude only 2.2% of the eukaryotic signal peptides in our data set), the percentage of false positives among the signal anchors drops to 28% for the eukaryotic and 32% for the human signal anchors (39% when using eukaryotic networks). When taking this into account, our method does provide a reasonably good discrimination between signal peptides and signal anchors. This has not been reported by any of the earlier published methods for signal peptide recognition.

Scanning the *Haemophilus influenzae* genome

We have applied the prediction method with networks trained on the Gram-negative data set to all the amino acid sequences of the predicted coding regions in the *Haemophilus influenzae* genome. The distribution of the mean S-score (from position 1 to the position with a maximal Y-score) is shown in Figure 4.

When applying the optimal cut-off value found for the Gram-negative data set, we obtained a crude estimate of the number of sequences with cleavable signal peptides in *H.influenzae*: 330 out of 1680 sequences or approximately 20%. If the maximal S-score is used instead of the mean S-score, the estimate comes out as 28% and with the maximal Y-score it is 14% (distributions not shown). If all three criteria

are applied together, leaving only 'typical' signal peptides, we obtain 188 sequences (11%).

Some of the sequences predicted to be signal peptides according to the S-score but not according to the Y-score may be signal anchor-like sequences of type II (single-spanning) or type IV (multispanning) membrane proteins. This hypothesis is strengthened by a hydrophobicity analysis of the ambiguous examples (results not shown). If we apply the slightly higher cut-off optimized for the discrimination of signal anchors versus signal peptides in eukaryotes (0.62) to the mean S-score, the estimate is lowered from 20 to 15%.

On the other hand, some of the sequences predicted to be signal peptides according to the maximal Y-score but not the mean S-score may be the effect of the initiation codon of the predicted coding region having been placed too far upstream. In this case, the apparent signal peptide becomes too long and the region between the false and the true initiation codon will probably not have signal peptide character, thereby bringing the mean S-score of the erroneously extended signal peptide region below the cut-off. This is strengthened by the finding that these ambiguous examples are longer than average and contain more methionines.

In conclusion, we estimate that 15–20% of the *H.influenzae* proteins are secretory. However, a whole-genome analysis like this would be more reliable if combined with other analyses, notably transmembrane segment predictions and initiation site predictions.

Method and data publicly available

The finished prediction method is available both via an e-mail server and a WWW server. Users may submit their own amino acid sequences in order to predict whether the sequence is a signal peptide and, if so, where it will be cleaved. We recommend that only the N-terminal part (say 50–70 amino acids) of the sequences is submitted, so that the interpretation of the output is not obscured by false positives further downstream in the protein.

The user is asked to choose between the network ensembles trained on data from Gram-positive, Gram-negative or eukaryotic organisms. We did not include the networks trained on the single-species data sets in the servers, since these did not improve the performance.

The values of the C-, S- and Y-scores are returned for every position in the submitted sequence. In addition, the maximal Y-score, maximal S-score and mean S-score values are given for the entire sequence and compared with the appropriate cut-offs. If the sequence is predicted to be a signal peptide, the position with the maximal Y-score is mentioned as the most likely cleavage site. A graphical plot in postscript format, similar to those in Figure 2, may be requested from the servers. We strongly recommend that a graphical plot is always used for the interpretation of the output. The plot may give hints about, for example, multiple cleavage sites or erroneously assigned initiation, which would not be found when using only the maximal or mean score values.

The address of the mail server is signalp@cbs.dtu.dk. For detailed instructions, send a mail containing the word 'help' only. The WWW server is accessible via the Center for Biological Sequence Analysis homepage at <http://www.cbs.dtu.dk/>.

All the data sets mentioned in Table I are available from an FTP server at <ftp://virus.cbs.dtu.dk/pub/signalp>. Retrieve the file README for detailed descriptions of the data and the format.

THIS PAGE BLANK (USPTO)

The FTP server and the mail server can both be accessed directly from the WWW server.

References

- Arrigo,P., Giuliano,F., Scalia,F., Rapallo,A. and Damiani,G. (1991) *CABIOS*, 7, 353-357.
- Bairoch,A. and Boeckmann,B. (1994) *Nucleic Acids Res.*, 22, 3578-3580.
- Brunak,S., Engelbrecht,J. and Knudsen,S. (1991) *J. Mol. Biol.*, 220, 49-65.
- Fleischmann,R. *et al.* (1995) *Science*, 269, 449-604.
- Gierasch,L.M. (1989) *Biochemistry*, 28, 923-930.
- Hirst,J.D. and Sternberg,M.J.E. (1992) *Biochemistry*, 31, 7211-7218.
- Hobohm,U., Scharf,M., Schneider,R. and Sander,C. (1992) *Protein Sci.*, 1, 409-417.
- Ladunga,L., Czakó,F., Csabai,I. and Geszti,T. (1991) *CABIOS*, 7, 485-487.
- Mathews,B. (1975) *Biochim. Biophys. Acta*, 405, 442-451.
- Nielsen,H., Engelbrecht,J., von Heijne,G. and Brunak,S. (1996a) *Proteins*, 24, 165-177.
- Nielsen,H., Engelbrecht,J., von Heijne,G. and Brunak,S. (1997) *Int. J. Neural Sys.*, in press.
- Presnell,S.R. and Cohen,F.E. (1993) *Annu. Rev. Biophys. Biomol. Struct.*, 22, 283-298.
- Qian,N. and Sejnowski,T.J. (1988) *J. Mol. Biol.*, 202, 865-884.
- Rapoport,T.A. (1992) *Science*, 258, 931-936.
- Rumelhart,D.E., Hinton,G.E. and Williams,R.J. (1986) In Rumelhart,D., McClelland,J. and the PDP Research Groups (eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1: Foundations*. MIT Press, Cambridge, MA, pp. 318-362.
- Schneider,G. and Wrede,P. (1993) *J. Mol. Evol.*, 36, 586-595.
- Schneider,T.D. and Stephens,R.M. (1990) *Nucleic Acids Res.*, 18, 6097-6100.
- Shannon,C.E. (1948) *Bell System Technol. J.*, 27, 379-423, 623-656.
- von Heijne,G. (1983) *Eur. J. Biochem.*, 133, 17-21.
- von Heijne,G. (1985) *J. Mol. Biol.*, 184, 99-105.
- von Heijne,G. (1986) *Nucleic Acids Res.*, 14, 4683-4690.
- von Heijne,G. (1988) *Biochim. Biophys. Acta*, 947, 307-333.
- von Heijne,G. (1990) *J. Membrane Biol.*, 115, 195-201.
- von Heijne,G. and Abrahmsén,L. (1989) *FEBS Lett.*, 244, 439-446.

Received April 19, 1996; revised September 2, 1996; accepted September 12, 1996

THIS PAGE BLANK (USPTO)

Signal Sequences The Limits of Variation

Gunnar von Heijne

Research Group for Theoretical Biophysics
Department of Theoretical Physics, Royal Institute of Technology
S-100 44 Stockholm, Sweden

(Received 29 October 1984, and in revised form 20 February 1985)

Variations in length and composition of the charged N-terminal, central hydrophobic and polar C-terminal regions in a large sample of signal sequences have been mapped, both as a function of the overall length of the sequence, and in an absolute sense, i.e. various "extremes" have been sought. The results show subtle differences between eukaryotic and prokaryotic sequences, but the general impression of signal sequences as being highly variable is reinforced. Criteria for a "minimal" signal sequence are suggested and discussed.

1. Introduction

In the process of protein export a central role is played by the signal sequence: an N-terminal segment that somehow initiates export whereupon it is cleaved from the mature protein. All signal sequences seem to be built along the same general lines, but the fine-structure of the design has only recently become a subject of study (von Heijne, 1983, 1984a,b; Perlman & Halvorson, 1983). Three structurally dissimilar regions have been recognized so far: a positively charged N-terminal region, a central hydrophobic region and a more polar C-terminal region that seems to define the cleavage site. These regions are present in all signal sequences, but the limits imposed upon them by the export machinery have not been systematically studied; in particular, it has not been ascertained whether they are all equally prone to variations in length and amino acid composition. This is an important question, since one of the outstanding features of the signal sequences taken as a group is their extraordinary variability in terms of overall length and amino acid sequence.

In this paper, eukaryotic and prokaryotic signal sequences are grouped according to their lengths, and the variations with length of the three structural regions (termed the n, h, and c-regions in what follows) are analysed. The analysis shows subtle differences between eukaryotic and prokaryotic sequences, and suggests "minimal" requirements that a fully functional signal sequence must conform to. Available data on non-functional mutant sequences, as well as on export-competent revertants, are discussed in the light of these requirements. Finally, the functional significance of

the results are assessed and related to current models of protein export.

2. Methods

The sample under study consists of 118 eukaryotic and 32 prokaryotic signal sequences, all with known cleavage sites. In the prokaryotic sample, no sequences known to be cleaved by the "lipoprotein signal peptidase" (Innis *et al.*, 1984) have been included. Unless otherwise indicated, references can be found in von Heijne (1984b).

Prokaryotic sequences: *Escherichia coli* maltose binding protein; phage pBR322 β -lactamase; phage M13 major and minor coat proteins; *E. coli* λ -receptor; *Salmonella typhimurium* histidine binding protein; *S. typhimurium* lysine-arginine-ornithine binding protein; *E. coli* leucine binding protein; *E. coli* leucine-isoleucine-valine binding protein; *E. coli* arabinose binding protein; *E. coli* galactose binding protein; *E. coli* chromosomal β -lactamase; *E. coli* ompA protein; *E. coli* ompF protein; *E. coli* ompC protein; *E. coli* lt enterotoxin A and B-subunits; *Bacillus subtilis* α -amylase; *E. coli* alkaline phosphatase; *E. coli* phoE protein; *Staphylococcus aureus* protein A; *Corynebacterium diphtheriae* toxin tox229; *E. coli* papA; phage I κ gene VII and gene III proteins; *Vibrio cholerae* toxins ctxA and ctxB (Mekalanos *et al.*, 1983); *E. coli* pilin k88; *E. coli* tolC (Hackett *et al.*, 1983); *E. coli* d-ribose binding protein (Groarke *et al.*, 1983); *Pseudomonas aeruginosa* exotoxin A (Gray *et al.*, 1984); *Pseudomonas* sp. carboxypeptidase G2 (Minton *et al.*, 1984).

Eukaryotic sequences: rat whey phosphoprotein; human serum albumin; rat α_1 -acid glycoprotein; mouse thyrotropin α -subunit; atlantic hagfish, anglerfish and human insulins; ovine β and κ -caseins; ovine α and β -lactalbumins; rabbit α -lactalbumin; mouse immunoglobulin H-chain (H-315); hybridoma immunoglobulin H-chain (93g7); rabbit immunoglobulin H-chain (P μ 3); mouse κ -immunoglobulin L-chains (L-41b, L-315 and

THIS PAGE BLANK (USPTO)

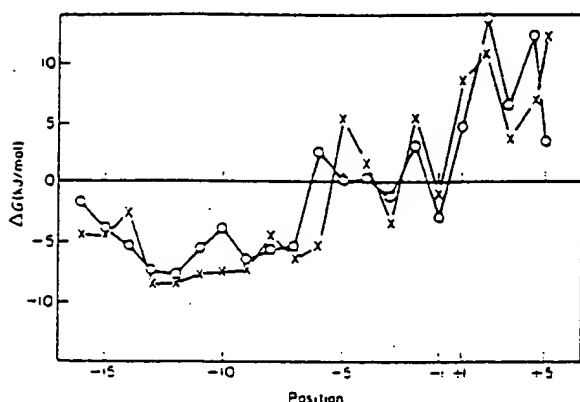


Figure 1. Mean hydrophobicity as a function of the position from the cleavage site (between -1 and +1) for the full eukaryotic (x) and prokaryotic (o) samples. The hydrophobicity scale used is from von Heijne (1981). Negative values are more hydrophobic.

L-321); mouse embryonic V_H -immunoglobulin (PCH 104); immunoglobulin V_H -107; caiman immunoglobulin V_H -gene; cockerel VLDL-II; bee melittin; rat lactin; human placental lactogen; human choriogonadotropin β -subunit; rabbit uteroglobin; rat growth hormone; bovine parathyroid hormone; human leukocyte, immune, γ and thymocyte interferons; rat relaxin; chicken tropoelastin B; chicken ovomucoid; chicken lysozyme; chicken albumin; human α_1 -antitrypsin; rat prostatic binding proteins C1, C2 and C3; mouse liver amylase; mouse MHC antigen H-2L₄; mouse MHC I-A light chain; mouse MHC I-E- α D chain; mouse MHC E₂ chain; human HLA-DR α and DR β_1 -chains; human HLA DC- α and DC- β chains; rat carboxypeptidase A; *Torpedo californica* acetylcholine receptor α , β , γ and δ -subunits; human muscle acetylcholine receptor; maize zein protein Z22.1; mouse C3 complement; rat pancreatic RNase; mouse opiomelanocortin; rat somatostatin; human antithrombin III; rat Thy-1; bovine Avp-NpII hormone; yeast invertase; mouse thyrotropin β -subunit; hamster glucagon; pea seed lectin; rat apolipoprotein E; barley α -amylase; human apolipoprotein A1; mouse β -crystallin; rat angiotensinogen; trypsinogen; rat elastase I and II; bovine chymosin; pea legumin; human insulin-like growth factor I; rat seminal vesicle secretion IV protein; rabbit poly-Ig receptor; VS virus glycoprotein G; adenovirus glycoprotein; rabies virus glycoprotein (ERA); human influenza A/Victoria/ and A/Jap/ haemagglutinins; avian influenza A/FPV/ haemagglutinin; *Herpes simplex* virus type-1 glycoprotein D; human parathyroid hormone (Hendy *et al.*, 1981); yeast pho5 (Arima *et al.*, 1983); *Herpes simplex* type 2 glycoprotein D gene (Watson, 1983); rat α -lactalbumin (Qasba & Safaya, 1984); bovine Ot-NpI hormone

(Ruppert *et al.*, 1984); human epidermal growth factor receptor (Ullrich *et al.*, 1984); human Christmas factor (Anson *et al.*, 1984); human pancreatic polypeptide (Boel *et al.*, 1984); wheat gliadin (Rafalski *et al.*, 1984); T-cell receptor α -subunit (Saito *et al.*, 1984); human transferrin (Yang *et al.*, 1984); hen ovotransferrin (Williams *et al.*, 1982); *Thaumatococcus daniellii* thaumatin II (Edens *et al.*, 1984); *Chironomus thummi thummi* globin IV (Antoine & Niessing, 1984); human atrial natriuretic factor (Nakayama *et al.*, 1984); human HTLV-I and HTLV-II envelope glycoproteins (Sodroski *et al.*, 1984); human insulin-like growth factor II (Dull *et al.*, 1984); murine epidermal growth factor binding protein (Lundgren *et al.*, 1984); calf acetylcholine receptor γ -subunit (Takai *et al.*, 1984); human α -fibrinogen (Kant *et al.*, 1983); human β and γ -fibrinogen (Chung *et al.*, 1983); human α -haptoglobin (Yang *et al.*, 1983); human retinol binding protein (Colantuoni, 1983); rat chymotrypsin (MacDonald *et al.*, 1982); *Drosophila melanogaster* glue protein 8 (Gartinkel, 1983); VS virus (N. J. Ogden) glycoprotein (Gallione, 1983); *Aplysia* R3-14 neuropeptide (Scheller *et al.*, 1984).

3. Results

(a) The length distribution is different for eukaryotic and prokaryotic signal sequences.

The number of sequences in the various length-classes is given in Table 1. Except for an extremely short sequence with length $L = 13$, the eukaryotic distribution starts at $L = 15$ (5 sequences) and the prokaryotic at $L = 18$ (3 sequences). Moreover, the main weight of the distributions falls between $L = 18$ and $L = 20$ for the eukaryotes (37%) and between $L = 21$ and $L = 23$ for the prokaryotes (59%). Thus, towards their lower ends the two distributions differ consistently by three residues.

(b) The mean position of the boundary between the h and c-regions is different in eukaryotes and prokaryotes and does not vary with overall length.

It has been noted that the overall amino acid composition of the c-region is more polar than that of the h-region (von Heijne, 1983). Indeed, in a plot of the mean hydrophobicity of each position in a large enough sample of signal sequences aligned from their cleavage sites, the h/c boundary stands out clearly (Fig. 1). There is an obvious difference between eukaryotes and prokaryotes, however: in eukaryotes, the mean position of the h/c boundary is between residues -6 and -5, whereas in prokaryotes it is between residues -7 and -6. This is true also for the individual length-classes defined

Number of residues

Position -10 -9

A. Eukaryotes

Leu	58	52
Val	11	11
Phe	8	6
Ala	11	18
Gly	2	2
Pro	2	0

B. Prokaryotes

Leu	5	11
Val	3	1
Phe	3	2
Ala	11	4
Gly	3	2
Pro	2	1

The h/c boundary -6 and -5 in the eukaryotes.

in Table 1 (data the boundary d. Apart from this each other close indicating that n but also parts similar selective

(c) The h-region residues.

The hydrophobicity and Trp are eukaryotic and sharply in frequency. Conversely, the Asp, Glu, Arg, Thr and Tyr are abundant in the h-region but do not vary appreciably. It has been noted that the amino acids in the h-region are more polar than those in the c-region (von Heijne, 1983). This is true also for the individual length-classes defined

Table 1

Number of signal sequences of given length in the eukaryotic and prokaryotic samples

Length	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	>30
Eukaryotes	1	0	(5	8	4)	13	19	12	9	(3	6)	14	10	(6	1	0	3	2)	2
Prokaryotes	0	0	0	0	0	(3	2	0	7)	(6	6)	(0	4	1	0	0	1)	0	2

The pooled length-classes referred to in the text are enclosed by parentheses when they encompass sequences of more than a single length.

THIS PAGE BLANK (USPTO)

Table 2

Number of residues in the h/c boundary region for a few selected amino acids

Position	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1
A. Eukaryotes											
Leu	58	52	34	43	29	6	19	5	17	7	6
Val	11	11	9	11	17	2	8	21	2	0	7
Phe	8	6	7	13	13	2	4	2	10	0	4
Ala	11	18	15	12	13	15	15	34	4	53	12
Gly	2	2	8	3	2	15	27	2	4	35	7
Pro	2	0	2	2	6	20	10	1	0	0	1
B. Prokaryotes											
Leu	5	11	12	2	0	2	3	1	2	0	0
Val	3	1	5	8	0	0	4	2	0	0	2
Phe	3	2	1	7	1	3	1	0	5	0	0
Ala	11	4	6	5	6	8	5	18	0	28	17
Gly	3	2	5	1	2	4	1	0	1	2	1
Pro	2	1	0	1	4	4	2	0	0	0	0

The h/c boundary (as defined in Fig. 1) is between positions -6 and -5 in the eukaryotes, and between -7 and -6 in the prokaryotes.

Table 3

Number of eukaryotic sequences with a given number of Leu, Phe and Gly residues in the region -13 to -6, and the expected numbers calculated on the assumption of a binomial distribution

No. of residues	0	1	2	3	4	5	6	7	8
No. of sequences									
Observed (Leu)	1	8	25	39	34	7	3	1	0
Expected (Leu)	2	11	26	33	27	14	5	1	0
Observed (Phe)	61	39	15	3	0	0	0	0	0
Expected (Phe)	59	43	13	2	0	0	0	0	0
Observed (Gly)	98	19	0	1	0	0	0	0	0
Expected (Gly)	95	21	2	0	0	0	0	0	0

(eukaryotes) or -15 to -7 (prokaryotes) with the numbers obtained for a sample with randomized h-regions (i.e. the amino acids in each individual h-region in the original sample were randomly "scrambled" before the pair-count was performed). As judged by χ^2 -analysis (one degree of freedom), no significant deviations ($P < 0.05$) from the expected counts were found, either for nearest-neighbours or for pairs separated by up to three residues, except for Leu-Ile (i, i+3)-pairs, which are about twice as numerous as expected (32 versus 13) in the eukaryotic sample (data not shown). This indicates that there are no strong sequence constraints in the h-region beyond the observed enrichment in hydrophobic residues.

(c) The h-region is enriched in hydrophobic residues, but has no apparent internal sequence regularities

The hydrophobic residues Phe, Ile, Leu, Met, Val and Trp are enriched in the h-regions of both eukaryotic and prokaryotic sequences, and drop sharply in frequency at the h/c boundary. Conversely, the charged and polar amino acids (Asp, Glu, Arg, Lys, His, Gly, Pro, Gln, Asn, Ser, Thr and Tyr) are virtually absent in the h-region but dominate the c-region. Ala, which is very abundant in the prokaryotic signal sequences, does not vary appreciably in incidence across the h/c boundary in either sample (Table 2).

It has been claimed that the distribution of amino acids in the h-region is non-random (Inouye & Haleboua, 1980; Perlman & Halvorson, 1983). In the present sample, however, no convincing patterns of fine-structure are apparent in this region: indeed, the number of eukaryotic sequences with a given number of a particular amino acid in the region -13 to -6 closely follows a random expectation, i.e. a binomial distribution (Table 3).

It has also been claimed that some nearest-neighbour pairs of amino acids in the h-region are present in numbers that cannot be explained on the basis of random pairing (Perlman & Halvorson, 1983). This possibility was tested in the present sample by comparing the observed number of pairs of given amino acids in the region -13 to -6

(d) The n-region accounts for one half of the length variation, but the net N-terminal charge does not vary with length

The net charge distribution in the n-region differs by one positive charge between eukaryotes and prokaryotes, indicating that the N-terminal amino group in eukaryotes provides one positive charge, whereas the blocked Met in prokaryotes does not (von Heijne, 1984b). As is clear from Figure 2(a), the net N-terminal charge does not vary appreciably with the overall length in either eukaryotes or prokaryotes, and has a mean value of about +1.7 in both groups. The length of the polar n-region varies strongly with the overall length, however (Fig. 2(b)): the variation is similar in eukaryotes and prokaryotes, and accounts for approximately one half of the total length variation.

(e) The h-region accounts for one half of the variation in overall length, but there are no regular variations in amino acid composition with length

Since the length of the c-region is independent of the total length, the remaining half of the length variation stems from the h-region. As is shown in Table 4, there are no regular variations in amino acid frequencies between the different length-classes, and the only suggestive observation so far is

THIS PAGE BLANK (USPTO)

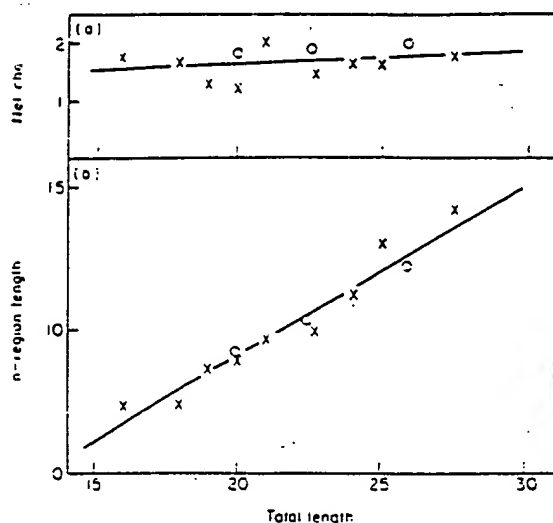


Figure 2. (a) Net charge and (b) length of the n-region as a function of the overall length for the length-classes defined in Table 1 (o, prokaryote; x, eukaryote). The initiator Met is assumed to provide one positive charge in the eukaryotes but not in the prokaryotes (cf. the text). The n/h boundary is defined by either (1) the last charged (Asp, Glu, Arg, Lys) or large polar (Asn, Gln, His, Pro) residues (whichever yields the longest n-region) on the N-terminal side of the uninterrupted non-polar region.

that the shortest sequences ($L=15$ in the eukaryotes, and $L=18$ in the prokaryotes) have the most hydrophobic h-regions (per residue) both in the eukaryotes and the prokaryotes (the mean hydrophobicity per residue in the region -13 to -6 in the eukaryotic $L=15$ sequences is -8.9 kJ/mol, the next-largest value is -8.3 kJ/mol (for the $L=30$ sequences); the value for the region -15 to -7 in the prokaryotic $L=18$ sequences is -9.6 kJ/mol, the next-largest value is -6.2 kJ/mol (for the $L=23$ sequences)).

(f) Examples of "extreme" sequences

A selection of sequences (extracted from our full collection of some 300 entries) that are "extreme" in one way or another is on display in Figure 3. The

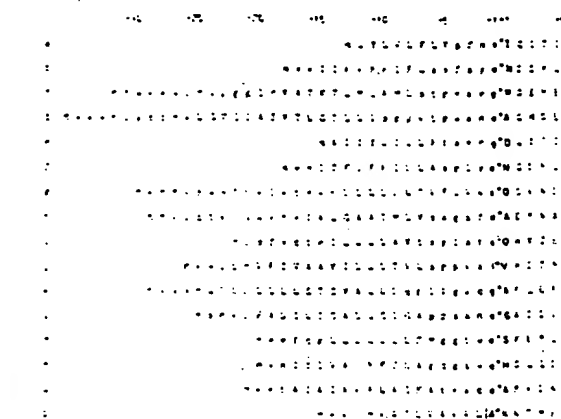


Figure 3. A collection of "extreme" signal sequences (see the text). The h-region is in boldface. Line a, mouse β -crystallin; b, *V. cholerae* toxin (ctxA); c, human HLA-DX β -chain; d, *S. aureus* protein A; e, human influenza A/Jap/ haemagglutinin; f, *E. coli* lt enterotoxin, A-subunit; g, human β fibrinogen; h, phage lke gene VIII-protein; i, chicken α -2 (I) collagen (Tate *et al.*, 1983); j, mouse MHC E β -chain; k, human pancreatic polypeptide; l, *C. diphtheriae* toxin tox 223; m, human chorionic gonadotropin, β -subunit; n, hamster glucagon; o, *E. coli* ompA protein; p, ovine α -S2 casein (Mercier & Gaye, 1980).

first four entries show extremes in overall length, with a more than twofold increase from the shortest ($L=13$) to the longest ($L=36$). The variation in the n-region is even more impressive: from one residue (line e) to 17 (line g). The shortest eukaryotic h-region found so far is only seven residues long (line i); the longest is some 16 residues (line k). For the prokaryotes, the corresponding values are nine and 15 (lines b and l). In terms of amino acid composition, there are h-regions that are almost 100% Leu (line m), 0% Leu (line n), rich in Phe (line b), and rich in Ala (line o). The sequence in line d, finally, has an unusually long c-region (10 residues).

(g) Point mutations, deletions and revertants

Counting from the cleavage site, almost all export-deficient point mutations described so far

Table 4
Frequencies of a few selected amino acid residues in the h-region of the various length-classes (cf. Table 1)

Length	15-17	18	19	20	21	22-23	24	25	26-30
Leu	0.37	0.39	0.35	0.36	0.43	0.32	0.36	0.28	0.36
Val	0.08	0.13	0.12	0.08	0.10	0.11	0.09	0.16	0.08
Phe	0.08	0.09	0.09	0.12	0.08	0.08	0.07	0.12	0.04
Ala	0.13	0.10	0.09	0.08	0.12	0.10	0.13	0.11	0.09
Ser	0.07	0.04	0.09	0.05	0.05	0.06	0.07	0.04	0.10

The h-region is defined as the region between position -6 and the mean position of the n/h boundary as read from Fig. 2(b).

Figure 4. Signal deletions (Silhavy boldface).

fall in the region Leu₋₁₇ → Pro mutations and revertants are shown that an intact h-region -14 (counting from the proper functional Gly₋₁) is the only indication in the h-region.

Taken at face value, this suggests that the h-region makes up the signal selective pressure. The c-region does not vary in overall length: it is from residue -1 to -13 (in the prokaryotes), and defining the cleavage site.

The n-region, however, is variable both in length and composition, but is positively correlated with a more hydrophobic h-region.

The h-region, however, is an enigma. The sequences and their lengths agree that residues -6 to -13 (in the eukaryotes) and constitute the h-region. Overall, there is no governing principle for substitutions of hydrophobic residues in h-regions that are enriched in a "non-comfortable" trypsin cleavage site.

THIS PAGE BLANK (USPTO)

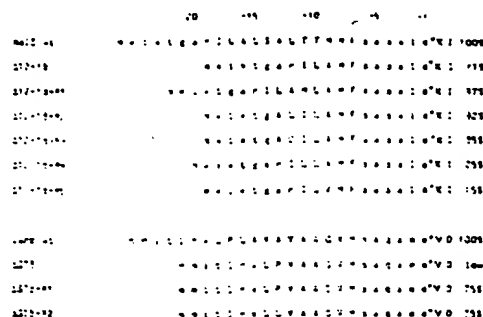


Figure 5. Signal sequence deletions and wholly or partially restored revertants (Bankaitis *et al.*, 1984; Emr & Silhavy, 1983). The h-region is in boldface. The percentage values refer to the amount of correctly exported protein relative to wild-type (wt).

mutations do not fit this picture without additional assumptions.

Thus, we are left with a picture of the "minimal" signal sequence as one composed of a five (eukaryotes) or six (prokaryotes) residue long c-region; a seven (eukaryotes) or eight (prokaryotes) residue long h-region with at most one Ser, Gly, Thr or Pro among the hydrophobic residues; and a one (eukaryotes) or two (prokaryotes) residue long, positively charged n-region. Thus, it seems that all three regions can be one residue shorter in eukaryotes, making the shortest eukaryotic sequences three residues shorter than the shortest prokaryotic ones (cf. Table 1). If the n, h and c-regions are indeed independent, it should be possible to make a functional 13-residue long eukaryotic signal sequence (cf. Fig. 3, line a) and a 16-residue long prokaryotic one. If the regions are allowed to overlap slightly, even shorter sequences may be possible (see Fig. 3, line p for an example of h/c overlap).

The maximal limits are harder to find. It seems clear that the c-region cannot be much longer than its "consensus" length of five or six residues. Likewise, if the h-region becomes longer than about 20 residues it may in fact anchor the protein permanently to the membrane and turn into an N-terminal *trans*-membrane sequence (cf. von Heijne, 1981; Bos *et al.*, 1984). The limits on the n-region, finally: 18 residues have been attached to the N terminus of a cloned insulin gene, making the n-region 21 residues long, with no effect upon export (Talmadge *et al.*, 1981); on the other hand, a mutant Sindbis virus glycoprotein with its N terminus fused to a 30,000 *M*_r cytoplasmic protein is not exported (Wirth *et al.*, 1979). The exact limit remains unknown, however.

What, then, are the functions of these regions? Structurally, they appear to be quite independent, and they do not seem to be co-selected in any important way. The c-region is clearly involved in defining the cleavage site (von Heijne, 1984a); the h-region has been suggested as being the target for

Taken at face value, the results presented here suggest that the n, h and c-regions that together make up the signal sequence are under different selective pressures, although they are by and large similarly selected in eukaryotes and prokaryotes. The c-region does not contribute to the variation in overall length; it extends, with small variations, from residue -1 to -5 (in eukaryotes) or -6 (in prokaryotes), and it follows the "(-3, -1)-rule" for defining the cleavage site (von Heijne, 1983, 1984a).

The n-region, on the other hand, is extremely variable both in terms of length and amino acid composition, but its net charge, which is always positive with a mean value of about +1.7, does not vary appreciably with its length.

The h-region, finally, still presents something of an enigma. The statistics on the wild-type sequences and the results from mutation studies agree that residues -7 to -14 (in prokaryotes) or -6 to -13 (in eukaryotes) are the most important ones and constitute what seems to be a "minimal" h-region. Overall hydrophobicity seems to be the one governing principle in this region, and, indeed, substitutions of more hydrophobic for less hydrophobic residues seem to make a big difference in h-regions that are close to the minimal length (Fig. 5): one Ser, Gly, Thr or Pro can obviously be tolerated in a "minimal" h-region, but not two. It is uncomfortably true, however, that the λ B Gly₋₉

THIS PAGE BLANK (USPTO)

signal recognition particle" (SRP) which is a translational block on cytoplasmic ribosomes synthesizing proteins destined for export (Walter *et al.*, 1981); and the n-region may have something to do with the "docking protein" catalysed release of the SRP-induced block (Hall *et al.*, 1983; Vlasuk *et al.*, 1983). In contrast to the universally conserved "(-3, -1)-design" of the cleavage site (von Heijne, 1983), however, no specific patterns of amino acids have yet been detected in the rest of the signal sequence, and the h as well as the n-region seem far too variable, both in terms of length and amino acid sequence, to allow for specific protein-protein interactions such as those envisioned in the current models of how SRP works. Instead, both regions seem well-suited for binding in a rather unspecific way to the surface (n-region) and to the interior (h-region) of membranes, as suggested repeatedly in the literature (DiRienzo *et al.*, 1978; von Heijne & Blomberg, 1979; Wickner, 1980; Engelman & Steitz, 1981).

The length variations observed in the h-region (~8 to ~20 residues) may indicate that this region spans the 25 to 30 Å thick non-polar interior of the membrane as a structure composed partly of α -helix, partly of extended chain depending on its length: eight residues in a fully extended structure will have a length of some 27 Å, close to the length of a 20-residue long helix. Since the helical conformation should be thermodynamically preferred in a non-polar environment, h-regions of intermediate length would span the membrane as part α , part extended-chain structures (cf. Bedouelle & Hofnung, 1981).

However, this does not explain the SRP-effect. By substituting a polar leucine analogue for Leu in a Leu-rich signal sequence one can bypass the SRP-induced translational block and get a cytoplasmic protein; with a Leu-poor signal sequence this does not happen (Walter *et al.*, 1981). It may be, though, that this does not come about as a result of a direct interaction between the signal sequence and the SRP; rather, the relation between the two may be indirect and not dependent upon a specific protein-protein interaction. One might speculate that the signal sequence somehow interacts with the ribosome rather than with the SRP, and that the SRP halts translation by interacting with the ribosomal translocation site on ribosomes "sensitized" by the presence of a signal sequence.

This work was supported by a grant from the Swedish Natural Sciences Research Council.

References

- Amara, D. S., Choo, K. H., Rees, D. J. G., Giannelli, F., Gould, K., Huddleston, J. A. & Brownlee, G. G. (1984). *EMBO J.* 3, 1053-1060.
- Bedouelle, H. & Niessing, J. (1984). *Nature (London)*, 310, 795-798.
- Arima, K., Oshima, T., Kubota, I., Nakamura, N., Mizunaga, T. & Toh-e, A. (1983). *Nucl. Acids Res.* 11, 1657-1672.
- Bankaitis, V. A., Rasmussen, B. A. & Bassford, P. J. (1984). *Cell*, 37, 243-252.
- Bedouelle, H. & Hofnung, M. (1981). In *Membrane Transport and Neuroreceptors*, pp. 399-403. Alan R. Liss, New York.
- Boel, E., Schwartz, T. W., Norris, K. E. & Fill, N. P. (1984). *EMBO J.* 3, 909-912.
- Bos, T. J., Davis, A. R. & Nayak, D. P. (1984). *Proc. Nat. Acad. Sci., U.S.A.* 81, 2327-2331.
- Chung, D. W., Chan, W.-Y. & Davie, E. W. (1983). *Biochemistry*, 22, 3250-3256.
- Colantuoni, V. (1983). *Nucl. Acids Res.* 11, 7769-7776.
- DiRienzo, J. M., Nakamura, K. & Inouye, M. (1978). *Annu. Rev. Biochem.* 47, 481-532.
- Dull, T. J., Gray, A., Hayflick, J. S. & Ullrich, A. (1984). *Nature (London)*, 310, 777-781.
- Edens, L., Bom, L., Ledeboer, A. M., Maat, J., Toonen, M. Y., Visser, C. & Verrips, C. T. (1984). *Cell*, 37, 629-633.
- Emr, S. D. & Silhavy, T. J. (1983). *Proc. Nat. Acad. Sci., U.S.A.* 80, 4599-4603.
- Engelman, D. M. & Steitz, T. A. (1981). *Cell*, 23, 411-422.
- Gallione, C. J. (1983). *J. Virol.* 46, 162-169.
- Garfinkel, M. D. (1983). *J. Mol. Biol.* 168, 765-789.
- Gray, G. L., Smith, D. H., Baldrige, J. S., Harkins, R. N., Vasil, M. L., Chen, E. Y. & Heyneker, H. L. (1984). *Proc. Nat. Acad. Sci., U.S.A.* 81, 2645-2649.
- Groarke, J. M., Mahoney, W. C., Hope, J. N., Furlong, C. E., Robb, F. T., Zalkin, H. & Hermodson, M. A. (1983). *J. Biol. Chem.* 258, 12952-12956.
- Hackett, J., Misra, R. & Reeves, P. (1983). *FEBS Letters*, 156, 307-309.
- Hall, M. N., Gabay, J. & Schwartz, M. (1983). *EMBO J.* 2, 15-19.
- Hendy, G. N., Kronenberg, H. M., Potts, J. T. & Rich, A. (1981). *Proc. Nat. Acad. Sci., U.S.A.* 78, 7365-7369.
- Innis, M. A., Tokunaga, M., Williams, M. E., Loranger, J. M., Chang, S.-Y., Chang, S. & Wu, H. C. (1984). *Proc. Nat. Acad. Sci., U.S.A.* 81, 3708-3712.
- Inouye, M. & Halegoua, S. (1980). *CRC Crit. Rev. Biochem.* 7, 339-371.
- Kant, J. A., Lord, S. T. & Crabtree, G. R. (1983). *Proc. Nat. Acad. Sci., U.S.A.* 80, 3953-3957.
- Lundgren, S., Ronne, H., Rask, L. & Peterson, P. A. (1984). *J. Biol. Chem.* 259, 7780-7784.
- MacDonald, R. J., Starv, S. J. & Swift, G. H. (1982). *J. Biol. Chem.* 257, 9724-9732.
- Mekalanos, J. J., Swartz, D. J., Pearson, G. D. N., Harford, N., Groyne, F. & de Wilde, M. (1983). *Nature (London)*, 306, 551-557.
- Mercier, J.-P. & Gaye, P. (1980). *Ann. N.Y. Acad. Sci.* 343, 232-251.
- Minton, N. P., Atkinson, T., Bruton, C. J. & Sherwood, R. F. (1984). *Gene*, 31, 31-38.
- Nakayama, K., Ohkubo, H., Hirose, T., Inayama, S. & Nakanishi, S. (1984). *Nature (London)*, 310, 699-707.
- Perlman, D. & Halvorson, H. O. (1983). *J. Mol. Biol.* 167, 391-409.
- Qasbi, P. K. & Safaya, S. K. (1984). *Nature (London)*, 308, 377-380.
- Rafalski, J. A., Scheets, K., Metzler, M., Peterson, D. M., Hedgroth, C. & Söll, D. G. (1984). *EMBO J.* 3, 1409-1415.
- Ruppert, S., Scherer, G. & Schutz, G. (1984). *Nature (London)*, 308, 554-557.
- Saito, H., Kranz, D., Eisen, H. N. & (London), 309, 7.
- Scheller, R. H., Kuhl, A. C., Nambu, J. R. & Science, 225, 1300.
- Silhavy, T. J., Ben-*Microbiol. Rev.* 47.
- Sodroski, J., Patarca, T.-H., Essex, M., R. C. & Haseltine, 424.
- Takai, T., Noda, M., Notake, M., Shirai, Tanaka, K., Hirai (1984). *Eur. J. Bi*
- Talmadge, K., Brosius (London), 294, 1.
- Tate, V. E., Finer, (1983). *Nucl. Ac*
- Ullrich, A., Coussens, Gray, A., Tam, Libermann, T. A., Mayes, E. L., V., Seeburg, P. H., 425.

THIS PAGE BLANK (USPTO)

- Nucl. Acids Res.* 11, 7769-7776.
- Inouye, M. (1978). *Membranes* 399-403. Alan R. Liss, Inc., New York.
- von Heijne, G. (1981). *Proc. Natl. Acad. Sci. U.S.A.* 78, 2264-2268.
- von Heijne, G. (1983). *Nature (London)* 309, 757-762.
- Scheller, R. H., Kaldany, R.-R., Kreiner, T., Mahon, A. C., Namhu, J. R., Schaefer, M. & Taussig, R. (1984). *Science* 225, 1300-1308.
- Silhavy, T. J., Benson, S. A. & Emr, S. D. (1983). *Microbiol. Rev.* 47, 313-344.
- Salasinski, J., Patarca, R., Perkins, D., Briggs, D., Lee, T.-H., Essex, M., Coligan, J., Wong-Staal, F., Gallo, R. C. & Haseltine, W. A. (1984). *Science* 225, 421-424.
- Takai, T., Noda, M., Furutani, Y., Takahashi, H., Notake, M., Shimizu, S., Kayano, T., Tanabe, T., Tanaka, K., Hirose, T., Inayama, S. & Numa, S. (1984). *Eur. J. Biochem.* 143, 109-115.
- Talmadge, K., Brosius, J. & Gilbert, W. (1981). *Nature (London)* 294, 176-178.
- Tate, V. E., Finer, M. H., Boedtker, H. & Doty, P. (1983). *Nucl. Acids Res.* 11, 91-104.
- Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Downward, J., Mayes, E. L. V., Whittle, N., Waterfield, M. D. & Seeburg, P. H. (1984). *Nature (London)* 309, 418-425.
- Vlasuk, G. P., Inouye, S., Ito, H., Itakura, K. & Inouye, M. (1983). *J. Biol. Chem.* 258, 7141-7148.
- von Heijne, G. (1981). *Eur. J. Biochem.* 120, 275-278.
- von Heijne, G. (1983). *Eur. J. Biochem.* 133, 17-21.
- von Heijne, G. (1984a). *J. Mol. Biol.* 173, 243-251.
- von Heijne, G. (1984b). *EMBO J.* 3, 2315-2318.
- von Heijne, G. & Blomberg, C. (1979). *Eur. J. Biochem.* 97, 175-181.
- Walter, P., Ibrahim, I. & Blobel, G. (1981). *J. Cell Biol.* 91, 545-561.
- Watson, R. J. (1983). *Gene* 26, 307-312.
- Wickner, W. (1980). *Science* 210, 861-868.
- Williams, J., Elleman, T. C., Kingston, I. B., Wilkins, A. G. & Kuhn, K. A. (1982). *Eur. J. Biochem.* 122, 297-303.
- Wirth, D. F., Lodish, H. F. & Robbins, P. W. (1979). *J. Cell Biol.* 81, 154-162.
- Yang, F., Brune, J. L., Baldwin, W. D., Barnett, D. R. & Bowman, B. H. (1983). *Proc. Natl. Acad. Sci., U.S.A.* 80, 5875-5879.
- Yang, F., Lum, J. B., McGill, J. R., Moore, C. M., Naylor, S. L., van Bragt, P. H., Baldwin, W. D. & Bowman, B. H. (1984). *Proc. Natl. Acad. Sci., U.S.A.* 81, 2752-2756.

Edited by R. Huber

THIS PAGE BLANK (USPTO)

Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in prokaryotic and eukaryotic cells

Gunnar von Heijne

Research Group for Theoretical Biophysics, Department of Theoretical Physics, Royal Institute of Technology, S-100 44 Stockholm, Sweden

Communicated by K. Simons

A statistical analysis of the distribution of charged residues in the N-terminal region of 39 prokaryotic and 134 eukaryotic signal sequences reveals a remarkable similarity between the two samples, both in terms of net charge and in terms of the position of charged residues within the N-terminal region, and suggests that the formyl group on Met_i is not removed in prokaryotic signal sequences.

Key words: amino acid statistics/protein export/signal sequences

Introduction

Secreted proteins are generally synthesized with an N-terminal signal sequence, some 15–25 amino acids long, that somehow initiates the export process. Several functions seem to be encoded within this short peptide: in eukaryotes, it initiates the interaction between the ribosome and the so-called signal recognition particle (SRP) which arrests translation until an unoccupied export-site on the membrane is found (Walter *et al.*, 1981) [for a discussion of a possible prokaryotic SRP, see Kumamoto *et al.* (1984)]; it may influence the 'SRP-receptor'-catalyzed release of SRP from the ribosome (Gilmore and Blobel, 1983; Hall *et al.*, 1983; Silhavy *et al.*, 1983); and it contains the information needed for its removal from the mature protein.

These different functions most likely reside in at least partially non-overlapping regions of the signal sequence (Figure 1). A C-terminal region of about five or six residues seems to be involved in defining the site of cleavage between the signal sequence and the mature protein (von Heijne, 1983, 1984; Perlman and Halvorson, 1983); the hydrophobic core of the signal sequence has been implicated in binding to SRP (Walter *et al.*, 1981); and the net charge of the N-terminal region has been shown to influence both the level of translation and the efficiency of export (Hall *et al.*, 1983; Vlasuk *et al.*, 1983).

Although the functions performed by the signal sequence are understood in a general sense, it has proved difficult to correlate them with either the amino acid sequence or higher order structures. In part, this is because the primary sequence has been found to be remarkably variable, and a very large number of sequences have to be analyzed if any significant patterns of amino acids, beyond the overall tri-partite design alluded to above, are to be found. So far, only a few studies of this kind have been published (von Heijne, 1983, 1984; Perlman and Halvorson, 1983), dealing chiefly with the amino acid pattern around the cleavage site between the signal sequence and the mature protein.

Here, the statistical distribution of charged amino acids in the N-terminal region in a large sample of known signal se-

quences is analyzed for the first time. The results show that prokaryotic and eukaryotic signal sequences, although superficially not very similar in this region, nevertheless seem to be under comparable selective pressures, yet another instance of the apparently very close homology between the two export machineries. The observed sequence differences can possibly be attributed to the fact that prokaryotic proteins are made with a formylated and uncharged N-terminal methionine, whereas eukaryotic proteins are initiated by an unformylated, positively charged Met residue.

Results and Discussion

The number of signal sequences with a given N-terminal net charge (counting +1 for Arg and Lys, and -1 for Asp and Glu) is given in Table 1. The distribution peaks at +1 for the eukaryotic sample and at +2 for the prokaryotes. Interestingly, the relative frequencies of N termini with a given net charge in the two samples overlap remarkably well if one positive charge is added to the eukaryotic sequences, Figure 2.

Likewise, the length distributions for the charged N-terminal regions have similar shapes (including only sequences with at least one charged residue besides the initiator Met in the samples, since the delineation of the polar N-terminal region is somewhat arbitrary in sequences lacking charged residues), except that the eukaryotic distribution peaks at a length of two residues, whereas the prokaryotic one has its main weight at a length of three residues, Figure 3.

An obvious explanation for these overlaps is that the selective constraints in terms of net charge on the prokaryotic and eukaryotic signal sequence N termini in fact are very similar, and that the initiator methionine (which is formylated and hence uncharged in prokaryotes, but unmodified and hence carrying a positively charged free amino group in eukaryotes) is not deformylated and remains bound to the N termini of the signal sequences. In the prokaryotes, an additional positively charged residue would thus be required, making these sequences on average one residue longer.

It is perhaps a little surprising that the uncharged Met_i in the prokaryotes must be compensated for by an additional Arg or Lys, since for cytoplasmic proteins deformylation and removal of the N-terminal Met are quite fast and take place shortly after the N terminus emerges from the ribosome (Pine, 1969; Housman *et al.*, 1972) – most likely, interactions with the export machinery (such as binding to SRP or membrane) competes with these cleavages, and protects Met_i from the deformylase.

If, in addition to the net charge, the positions of the charged residues within the N-terminal region are important, and if it is true that Met_i is not deformylated, we might expect to see a difference in the distribution of the charged residues between prokaryotes and eukaryotes since the latter automatically have a positive charge right at the N terminus. To test this, the number of sequences with either Arg or Lys next to the initiator Met was counted, only including N-terminal regions

THIS PAGE BLANK (ISPTO)

⊕ MET ASP TYR TYR ARG LYS TYR ALA ALA ILE PHE LEU VAL THR LEU SER VAL PHE LEU HIS VAL LEU HIS SER ALA ...

Fig. 1. A 'typical' signal sequence (human chorionic gonadotropin, α subunit) with the hydrophobic core in boldface and the cleavage site marked by an arrow. Note the charged residues in the N-terminal region, which has a net charge of +2.

Table 1. Number of signal sequences with a given net charge in the N-terminal region (counting +1 for Arg and Lys, and -1 for Asp and Glu) in the eukaryotic and prokaryotic samples

	Net charge							
	-2	-1	0	+1	+2	+3	+4	+5
Eukaryotes	2	8	40	68	13	2	1	0
Prokaryotes	0	0	0	7	24	5	1	2

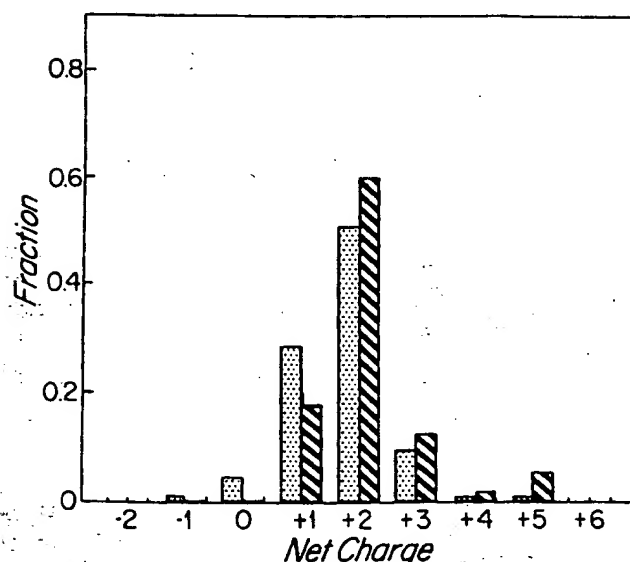


Fig. 2. Distribution of the net N-terminal charge calculated from Table 1, with one positive charge added to each eukaryotic sequence. Eukaryotic sample = dotted bars, prokaryotic sample = hatched bars.

(defined as the segment from the initiator Met to the charged residue closest to the hydrophobic core) longer than three residues (eukaryotes) or four residues (prokaryotes) in the calculation. Amongst the eukaryotic sequences, only six out of 62 (10%) have Arg or Lys in this position, whereas there are 12 sequences out of 22 (55%) with a positively charged residue next to Met_i among the prokaryotic sequences. Thus, a positive charge close to the N terminus may be advantageous.

Similarly, in signal sequences with both positively and negatively charged residues in their N-terminal region (such as the one in Figure 1), the charged region ends with a positive charge on its C-terminal side in 13 out of 14 cases. This latter observation may be explained either as a functional requirement or simply as a consequence of selection against negatively charged residues: starting from a highly positively charged N-terminal region (the only kind that presumably will accept a negative charge), an acidic residue should be much more likely to be functionally neutral when placed well inside the charged region rather than in the hydrophobic region, C-terminally to the last positive charge, where the introduction of a charged residue may often have a deleterious effect

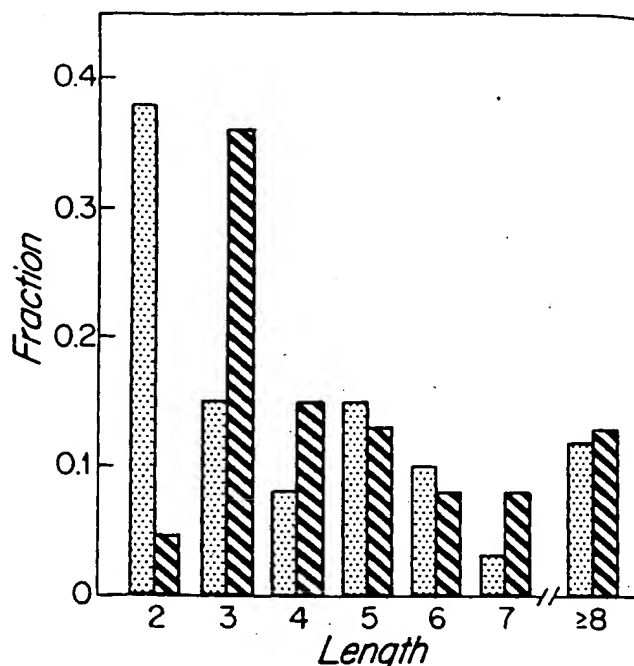


Fig. 3. Length distribution of the charged N-terminal regions (counting from the initiator Met to the charged residue closest to the hydrophobic core, see text). Eukaryotic sample = dotted bars, prokaryotic sample = hatched bars.

(Emr and Silhavy, 1982).

Although the connection between the selective constraints on the N-terminal region demonstrated here and the actual functioning of the signal sequence still remains to be made, one clue is provided by the work of Vlasuk *et al.* (1983) who have shown that both translation and secretion of the outer membrane lipoprotein from *Escherichia coli* are reduced when the net N-terminal charge is lowered from its wild-type value of +2: with zero net charge, synthesis is down to ~60% but secretion is unaffected; with a negative net charge, synthesis is reduced even further and, in addition, most of the protein made accumulates as precursor in the cytoplasm. This fits nicely with the net charge distribution in our sample, where 94% of all sequences have an N-terminal net charge of +1 or greater and 67% have a net charge of at least +2. There are only two sequences with a net charge less than zero, i.e., a α -immunoglobulin light chain (L-321) (Burstein and Schechter, 1978), and pea seed lectin (Higgins *et al.*, 1983); unfortunately, not much is known about their level of synthesis and secretion.

Finally, if the explanation given above for the difference in the number of basic residues between the prokaryotic and eukaryotic signal sequences (i.e., the uncharged Met_i versus the positively charged Met at their respective N termini) is correct, one prediction is that archaeobacterial signal sequences will turn out to be more like their eukaryotic than prokaryotic counterparts since they are synthesized with an unformylated initiator Met. In view of the results of Vlasuk *et*

THIS PAGE BLANK (USPTO)

al. referred to above, another implication is that up to 30% of all eukaryotic secreted proteins (those with a net charge of or less, not counting the initiator Met) may need one or extra positively charged residues on their N terminus to be effectively made and exported from a prokaryotic cell.

Materials and methods

Thirty nine prokaryotic and 134 eukaryotic signal sequences are included in this study. Unless otherwise indicated, the original references can be found in Michaelis and Beckwith (1982) or von Heijne (1983, 1984).

Prokaryotic sequences: the sample includes all sequences listed in von Heijne (1984), plus the following: *Bacillus licheniformis* and *Staphylococcus aureus* penicillinase (Nielsen and Lampen, 1982); *E. coli* and *Erwinia amylovora* lipoprotein (Yamagata *et al.*, 1981); *E. coli* heat stable toxin ST1; *Corynebacterium diphtheriae* toxin (Kaczorek *et al.*, 1983); *E. coli* galactose-binding protein (Scripture and Hogg, 1983); phage IKe major and minor coat proteins (P. Peeters and R. Konings, personal communication); *Vibrio cholerae* toxin (Mekalanos *et al.*, 1983); *E. coli* papA, papC, papE, papF, papG, papH, pilin K88, and pilin K88a (M. Baga, S. Lundberg, B. Lund and S. Normark, personal communication); and molA (Clément and Hofnung, 1981).

Eukaryotic sequences: the sample includes all sequences listed in von Heijne (1984), except: mouse embryonic V_H immunoglobulin; rat α immunoglobulin; human chorionic gonadotropin α subunit; mouse λ immunoglobulin; canine trypsinogen 2 + 3; human ϵ chain immunoglobulin; rabies virus glycoprotein CVS; and rat pituitary glycoprotein hormone α subunit. The sample also includes: murine and human α -fetoprotein (Law and Dugaiczky, 1981; Morinaga *et al.*, 1983); human leukocyte interferon (Gray and Goeddel, 1982; Taniguchi *et al.*, 1980); human V_H, V_H-107, and H-chain (93g7) immunoglobulin (Bernstein *et al.*, 1982; Early *et al.*, 1980; Sims *et al.*, 1982); porcine gastrin; human influenza haemagglutinin A/Jap/ and avian influenza haemagglutinin A/FPV/; mouse α immunoglobulins (L-41B, L-315, and L-321) (Burststein and Schechter, 1978); mouse embryonic V_H immunoglobulin; ovine α -S1 and α -S2 casein; mouse MHC H-2L_d and E_g chains (Evans *et al.*, 1982; Saito *et al.*, 1983); mouse MHC I-A light chain (Malissen *et al.*, 1983); mouse MHC I-E α -D chain (Hyldig-Nielsen *et al.*, 1983); mouse catfish and anglerfish somatostatin (Magazin *et al.*, 1982; Argos *et al.*, 1983); Goodman *et al.*, 1980; Hobart *et al.*, 1980); human and porcine ACTH- β -LPH precursor; trypanosome surface glycoprotein; anglerfish glucagon; flounder antifreeze protein (Davies *et al.*, 1982); herpes simplex virus glycoproteins C and D (Frink *et al.*, 1983; Watson *et al.*, 1982); mouse C3 complement (Wiebauer *et al.*, 1982); bovine opiomelanocortin (Uhlir and Herbert, 1983); human plasminogen activator (Pennica *et al.*, 1983); mouse epidermal growth factor (Gray *et al.*, 1983); pea seed lectin (Higgins *et al.*, 1983); rat apolipoprotein E (McLean *et al.*, 1983); rat prostatic binding protein C3 (Viskochil *et al.*, 1983); barley α -amylase (Rogers and Millman, 1983); human apolipoprotein A-I (Karathanasis *et al.*, 1983); human HLA-DC α chain (Schenning *et al.*, 1984); human HLA-DC β chain (Larhammar *et al.*, 1983); human HLA-DS α chain (Chang *et al.*, 1983); human vasoactive intestinal polypeptide (Itoh *et al.*, 1983); mouse β -crystalline (Inana *et al.*, 1983); human interleukin II (Taniguchi *et al.*, 1983); bovine kininogen (Nawa *et al.*, 1983b); french bean phaseolin (Slightom *et al.*, 1983); rat angiotensinogen (Doolittle, 1983); yeast mating factor α -1, trypsinogen, rat elastase I and II, bovine chymosin, fruit fly vitellogenin I, human and bovine lactin (Dayhoff, 1983); bovine substance P (Nawa *et al.*, 1983a); mouse interleukin-3 (Fung *et al.*, 1984); *Amaranthus hybridus* psbA (Hirschenberg and McIntosh, 1983); pea legumin (Lycett *et al.*, 1984); *Aplysia* egg-laying hormone (Scheller *et al.*, 1983); human insulin-like growth factor (Jansen *et al.*, 1983); yeast K1 toxin (Skipper *et al.*, 1984); Friend spleen focus-forming virus gp55 glycoprotein (Amanuma *et al.*, 1983); rat luteinizing hormone β -subunit (Chin *et al.*, 1983); rat seminal vesicle secretion IV protein (Harris *et al.*, 1983); human growth hormone-releasing factor (Gubler *et al.*, 1983); mouse T-cell receptor (Hedrick *et al.*, 1984); human T cell protein YT35 (Yanagi *et al.*, 1984); rabbit poly-Ig receptor (Mostov *et al.*, 1984); and rat cholecystokinin (Deschenes *et al.*, 1984).

Acknowledgements

This work was carried out during a stay at the Department of Microbiology and Molecular Genetics, Harvard Medical School; my thanks to Professor Michael Beckwith for his support and hospitality, and to Dr. Kathy Strauch who taught me a lot about the biochemistry of Met.

References

Amanuma, H., Katori, A., Obata, M., Sagata, N. and Ikawa, Y. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3913-3917.

- Argos, P., Taylor, W.L., Minth, C.D. and Dixon, J.E. (1983) *J. Biol. Chem.*, **258**, 8788-8793.
- Bernstein, K.E., Premkumar Reddy, E., Alexander, C.B. and Mage, R.G. (1982) *Nature*, **300**, 74-76.
- Burststein, Y. and Schechter, I. (1978) *Biochemistry (Wash.)*, **17**, 2392-2400.
- Chang, H.-C., Moriuchi, T. and Silver, J. (1983) *Nature*, **305**, 813-815.
- Chin, W.W., Godine, J.E., Klein, D.R., Chang, A.S., Tan, L.K. and Habener, J.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4649-4653.
- Clément, J.M. and Hofnung, M. (1981) *Cell*, **27**, 507-514.
- Davies, P.L., Roach, A.H. and Hew, C.-L. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 335-339.
- Dayhoff, M. (1983) *Protein Sequence Data Bank*.
- Deschenes, R.J., Lorentz, L.J., Haun, R.S., Roos, B.A., Collier, K.J. and Dixon, J.E. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 726-730.
- Doolittle, R.F. (1983) *Science (Wash.)*, **222**, 417-419.
- Early, P., Huang, H., Davis, M., Calame, K. and Hood, L. (1980) *Cell*, **19**, 981-992.
- Emr, S.D. and Silhavy, T.J. (1982) *J. Cell Biol.*, **95**, 689-696.
- Evans, G.A., Margulies, D.H., Camerini-Otero, R.D., Ozato, K. and Seidman, J.G. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 1994-1998.
- Frink, R.J., Eisenberg, R., Cohen, G. and Wagner, E.K. (1983) *J. Virol.*, **45**, 634-647.
- Fung, M.C., Hapel, A.J., Ymer, S., Cohen, D.R., Johnson, R.M., Campbell, H.D. and Young, I.G. (1984) *Nature*, **307**, 233-237.
- Furutani, Y., Morimoto, Y., Shibahara, S., Noda, M., Takahashi, H., Hirose, T., Asai, M., Inayama, S., Hayashida, H., Miyata, T. and Numa, S. (1983) *Nature*, **301**, 537-540.
- Gilmore, R. and Blobel, G. (1983) *Cell*, **35**, 677-685.
- Goodman, R.H., Jacobs, J.W., Chin, W.W., Lund, P.K., Dee, P.C. and Habener, J.F. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5869-5873.
- Gray, A., Dull, T. and Ullrich, A. (1983) *Nature*, **303**, 722-725.
- Gray, P.W. and Goeddel, D.V. (1982) *Nature*, **298**, 859-863.
- Gubler, U., Monahan, J.J., Lomedico, P.T., Bhatt, R.S., Collier, K.J., Hoffman, B.J., Böhlen, P., Esch, F., Ling, N., Zeytin, F., Brazeau, P., Poonian, M.S. and Gage, L.P. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4311-4314.
- Hall, M.N., Gabay, J. and Schwartz, M. (1983) *EMBO J.*, **2**, 15-19.
- Harris, S.E., Mansson, P.-E., Tully, D.B. and Burkhart, B. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6460-6464.
- Hedrick, S.M., Nielsen, E.A., Kavalier, J., Cohen, D.I. and Davis, M.M. (1984) *Nature*, **308**, 153-158.
- Higgins, T.J.V., Chandler, P.M., Zurawski, G., Button, S.C. and Spencer, D. (1983) *J. Biol. Chem.*, **258**, 9544-9549.
- Hirschenberg, J. and McIntosh, L. (1983) *Science (Wash.)*, **222**, 1346-1349.
- Hobart, P., Crawford, R., Shen, L., Pictet, R. and Rutter, W.J. (1980) *Nature*, **288**, 137-141.
- Housman, D., Gillespie, D. and Lodish, H.F. (1972) *J. Mol. Biol.*, **65**, 163-166.
- Hyldig-Nielsen, J.J., Schenning, L., Hammerling, U., Widmark, E., Heldin, E., Lind, P., Servenius, B., Lund, T., Flavell, R., Lee, J.S., Trowsdale, J., Schreier, P.H., Zablitsky, F., Larhammar, D., Peterson, P.A. and Rask, L. (1983) *Nucleic Acids Res.*, **11**, 5055-5071.
- Inana, G., Piatigorsky, J., Norman, B., Slingsby, C. and Blundell, T. (1983) *Nature*, **302**, 310-315.
- Itoh, N., Obata, K., Yanai, N. and Okamoto, H. (1983) *Nature*, **304**, 547-549.
- Jansen, M., van Schaik, F.M.A., Ricker, A.T., Bullock, B., Woods, D.E., Gabbay, K.H., Nussbaum, A.L., Sussenbach, K.H. and Van der Brande, J.L. (1983) *Nature*, **306**, 609-611.
- Kaczorek, M., Delpeyroux, F., Chenciner, N., Sreeck, R.E., Murphy, J.R., Boquet, P. and Tiollais, P. (1983) *Science (Wash.)*, **221**, 855-858.
- Karathanasis, S.K., Zannis, V.I. and Breslow, J.L. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6147-6151.
- Kumamoto, C.A., Oliver, D.B. and Beckwith, J. (1984) *Nature*, **308**, 863-864.
- Larhammar, D., Hyldig-Nielsen, J.J., Servenius, B., Andersson, G., Rask, L. and Peterson, P.A. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7313-7317.
- Law, S.W. and Dugaiczky, A. (1981) *Nature*, **291**, 201-205.
- Lycett, G.W., Croy, R.R.D., Shirsat, A.H. and Boulter, D. (1984) *Nucleic Acids Res.*, in press.
- Magazin, M., Minth, C.D., Funckes, C.L., Deschenes, R., Tavianini, M.A. and Dixon, J.E. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 5152-5156.
- Malissen, M., Hunkapiller, T. and Hood, L. (1983) *Science (Wash.)*, **221**, 750-754.
- McLean, J.W., Fukazawa, C. and Taylor, J.M. (1983) *J. Biol. Chem.*, **258**, 8993-9000.
- Mekalanos, J.J., Swartz, D.J., Pearson, G.D.N., Harford, N., Groyne, F. and de Wilde, M. (1983) *Nature*, **306**, 551-557.
- Michaelis, S. and Beckwith, J. (1982) *Annu. Rev. Microbiol.*, **36**, 435-465.
- Morinaga, T., Sakai, M., Wegmann, T.G. and Tamaoki, T. (1983) *Proc. Natl.*

THIS PAGE BLANK (USPTO)

- Acad. Sci. USA*, 80, 4604-4608.
- Mostov, K.E., Friedlander, M. and Blobel, G. (1984) *Nature*, 308, 37-43.
- Nawa, H., Hirose, T., Takashima, H., Inayama, S. and Nakanishi, S. (1983a) *Nature*, 306, 32-36.
- Nawa, H., Kitamura, N., Hirose, T., Asai, M., Inayama, S. and Nakanishi, S. (1983b) *Proc. Natl. Acad. Sci. USA*, 80, 90-94.
- Nielsen, J.B.K. and Lampen, J.O. (1982) *J. Biol. Chem.*, 257, 4490-4495.
- Pennica, D., Holmes, W.E., Kohr, W.J., Harkins, R.N., Vohar, G.A., Ward, C.A., Bennett, W.F., Yelverton, E., Seeburg, P.H., Heyneker, H.L. and Goeddel, D.V. (1983) *Nature*, 301, 214-221.
- Perlman, D. and Halvorson, H.O. (1983) *J. Mol. Biol.*, 167, 391-409.
- Pine, M.J. (1969) *Biochim. Biophys. Acta*, 174, 359-372.
- Rogers, J.C. and Millman, C. (1983) *J. Biol. Chem.*, 258, 8169-8174.
- Saito, H., Maki, R.A., Clayton, L.K. and Tonegawa, S. (1983) *Proc. Natl. Acad. Sci. USA*, 80, 5520-5524.
- Scheller, R.H., Jackson, J.F., McAllister, L.B., Rothman, B.S., Mayeri, E. and Axel, R. (1983) *Cell*, 32, 7-22.
- Schenning, L., Larhammar, D., Bill, P., Wiman, K., Jonsson, A.-K., Rask, L. and Peterson, P.A. (1984) *EMBO J.*, 3, 447-452.
- Scripture, J.B. and Hogg, R.W. (1983) *J. Biol. Chem.*, 258, 10853-10855.
- Silhavy, T.J., Benson, S.A. and Emr, S.D. (1983) *Microbiol. Rev.*, 47, 313-344.
- Sims, J., Rabbitts, T.H., Estess, P., Slaughter, C., Tucker, P.W. and Capra, J.D. (1982) *Science (Wash.)*, 216, 309-310.
- Skipper, N., Thomas, D.Y. and Lau, P.C.K. (1984) *EMBO J.*, 3, 107-111.
- Slightom, J.L., Sun, S.M. and Hall, T.C. (1983) *Proc. Natl. Acad. Sci. USA*, 80, 1897-1901.
- Taniguchi, T., Mantel, N., Schwarstein, M., Nagata, S., Muramatsu, M. and Weissmann, C. (1980) *Nature*, 285, 547-549.
- Taniguchi, T., Matsui, H., Fujita, T., Takaoka, C., Kashima, N., Yoshimoto, R. and Hamuro, J. (1983) *Nature*, 302, 305-310.
- Uhler, M. and Herbert, E. (1983) *J. Biol. Chem.*, 258, 257-261.
- Viskochil, D.H., Perry, S.T., Lea, O.A., Stafford, D.W., Wilson, E.M. and French, F.S. (1983) *J. Biol. Chem.*, 258, 8861-8866.
- Vlasuk, G.P., Inouye, S., Ito, H., Itakura, K. and Inouye, M. (1983) *J. Biol. Chem.*, 258, 7141-7148.
- von Heijne, G. (1983) *Eur. J. Biochem.*, 133, 17-21.
- von Heijne, G. (1984) *J. Mol. Biol.*, 173, 243-251.
- Walter, P., Ibrahim, I. and Blobel, G. (1981) *J. Cell Biol.*, 91, 545-550.
- Watson, R.J., Weis, J.H., Salstrom, J.S. and Enquist, L.W. (1982) *Science (Wash.)*, 218, 381-384.
- Wiebauer, K., Domdey, H., Diggelmann, H. and Fey, G. (1982) *Proc. Natl. Acad. Sci. USA*, 79, 7077-7081.
- Yamagata, H., Nakamura, K. and Inouye, M. (1981) *J. Biol. Chem.*, 256, 2194-2198.
- Yanagi, Y., Yoshikai, Y., Leggett, K., Clark, S.P., Aleksander, I. and Mak, T.W. (1984) *Nature*, 308, 145-149.

Received on 29 June 1984

THIS PAGE BLANK (USPTO)

Patterns of Amino Acids near Signal-Sequence Cleavage Sites

Gunnar VON HEIJNE

Research Group for Theoretical Biophysics, Department of Theoretical Physics, Royal Institute of Technology, Stockholm

(Received October 27/December 30, 1982) – EJB 6142

According to the signal hypothesis, a signal sequence, once having initiated export of a growing protein chain across the rough endoplasmic reticulum, is cleaved from the mature protein at a specific site. It has long been known that some part of the cleavage specificity resides in the last residue of the signal sequence, which invariably is one with a small, uncharged side-chain, but no further specific patterns of amino acids near the point of cleavage have been discovered so far. In this paper, some such patterns, based on a sample of 78 eukaryotic signal sequences, are presented and discussed, and a first attempt at formulating rules for the prediction of cleavage sites is made.

The molecular basis for the highly precise cleavage between signal sequences and mature exported proteins has been somewhat of an enigma ever since the first pre-proteins were sequenced some ten years ago [1]. The last residue of the signal sequence has been found to be either Ala, Gly, Ser, Cys, Thr or Gln, i.e. one with a small, uncharged side-chain, and it has also been shown that the signal sequence has a very hydrophobic central core [2], but neither the positioning of this core relative to the cleavage site, nor the patterns of amino acids found at this site have so far allowed a precise prediction of the position of the point of cleavage from sequence data alone. In this paper, we show that the patterning of amino acids near the cleavage site is much richer than hitherto believed, and that, given this information, cleavage sites can be predicted quite successfully.

RESULTS AND DISCUSSION

Analysis of a comprehensive collection of known eukaryotic signal sequences, aligned with coincident cleavage sites (Table 1), reveals some interesting regularities. As shown in Table 2, whole classes of residues are absent not only from position -1 (the last residue of the signal sequence, cf. above) but also from position -3 , and, conversely, the same types of amino acids as in position -1 are strongly preferred in position -3 . Also, these small, neutral residues are quite rare in position -2 (there is only one Ala in position -2 , in contrast to position -1 with 36, and position -3 with 25 alanines).

Moreover, the bulky aromatic residues seem to be selected for in position -2 , whereas the secondary-structure-disrupting residues Gly and Pro are found predominantly in positions -4 and -5 , respectively. In addition, there is not a single Pro residue in the whole region between -3 and $+1$.

The hydrophobic residues seem to be merely tolerated in positions $+1$ and -2 to -4 : they are conspicuously rare in position -5 , and then become dominant from position -6 onwards as one goes further into the central hydrophobic core (see above).

Finally, although charged residues are tolerated to some extent in positions -2 , -4 and -5 , they are much more abundant in the mature sequences, starting from position $+1$.

A similar analysis of a smaller sample with all isozymes and obviously related sequences removed yields essentially identical results (data not shown).

These observations make it possible to formulate a preliminary set of rules that can be used to predict the point of cleavage in a given pre-protein sequence. At this stage, we have felt it reasonable to deal with classes of amino acids rather than individual ones, and the following scheme can thus be refined as more sequences and data for individual residues become available.

The prediction method proposed here has two components: first, the N terminus of the hydrophobic core is located by searching from the N terminus for the first quadruplet with at least three hydrophobic residues (group IV in Table 2, not counting the initiator Met), and defining a 'window' for processing between residues $i+12$ and $i+20$ (i being the first residue in the quadruplet). Then, for each residue in the 'window', a measure of 'processing probability' is calculated by multiplying together the respective values from Table 3, values chosen 'by inspection' so as roughly to reflect the amino acid patterns discussed above, and fine-tuned to maximize the number of correct predictions. Finally, the site of processing is predicted as the site with the highest 'processing probability'.

Applying this scheme to the sequences listed in Table 1, only 5 out of 76 processing sites are incorrectly predicted (nos 18, 22, 37, 42, 45), with one additional case where the method cannot differentiate between the correct and one incorrect site (no. 52).

Using less than the full set of rules, it was found that the rule 'no Pro in position $+1$ ' can be discarded without affecting the number of correct and undecided predictions for the present data base. All other rules are important, though. As an illustration, using only the rule for position -1 yields 16 incorrect and 39 undecided predictions, and adding the rule for position -3 only improves these figures to 14 incorrect and 16 undecided predictions.

The amino acid patterns presented here can perhaps also yield some insight into the actual structure of the signal-sequence–protease complex. As shown above, positions -1 and -3 seem to be strongly selected for small, neutral residues, whereas positions $+1$, -2 , and -4 seem to accommodate almost any kind of residue. Moreover, there is not a single Pro from position -3 to $+1$, and, further, this region is often separated from the hydrophobic core of the signal sequence by the strong helix-breakers Pro or Gly [3].

Thus, a possible signal-sequence–protease structure would be an alternating sheet structure near the cleavage site, joined at its N terminus to a largely hydrophobic helix (Fig. 1). Possibly, position $+1$ is located close to the membrane surface, as

THIS PAGE BLANK (USPTO)

Table 1. A collection of eukaryotic signal sequences

In the first part, the sequences have been aligned from their known cleavage sites (between positions -1 and +1): in the second part sequences in which the location of this site is unknown have been aligned from their predicted sites of cleavage (see text). The predicted cleavage sites are indicated by a star (*), and the processing 'window' is shown by slashes (/); when a * and a / coincides, only the * is shown). Amino acids are symbolized by the one-letter code, i.e. A = Ala, C = Cys, D = Asp, E = Glu, F = Phe, G = Gly, H = His, I = Ile, K = Lys, L = Leu, M = Met, N = Asn, P = Pro, Q = Gln, R = Arg, S = Ser, T = Thr, V = Val, W = Trp, Y = Tyr, X = unknown. Proteins are as follows: 1: whey phosphoprotein, rat; 2: α -1 acid glycoprotein, rat; 3: α -thyrotropin, mouse; 4: insulin, hagfish; 5: insulin, anglerfish; 6: insulin, human; 7: insulin I, rat; 8: insulin II, rat; 9: β -casein, ovine; 10: α -casein, ovine; 11: α -lactalbumin, ovine; 12: β -lactoglobulin, ovine; 13: α -s1 casein, ovine; 14: α -s2 casein, ovine; 15: glycoprotein, VS virus; 16: VLDL-II, cockerel; 17: melittin, bee; 18: lactin, rat; 19: placental lactogen, human; 20: β -choriogonadotropin, human; 21: α -choriogonadotropin, human; 22: uteroglobin, rabbit; 23: growth hormone, rat; 24: growth hormone, human; 25: growth hormone, bovine; 26: parathyroid hormone, bovine; 27: relaxin, rat; 28: serum albumin, rat; 29: serum albumin, human; 30: liver albumin, rat; 31: tropoelastin B, chicken; 32: ovomucoid, chicken; 33: lysozyme, chicken; 34: conalbumin, chicken; 35: α -1 antitrypsin, human; 36: prostatic binding protein, rat; 37: prostatic binding protein c2, rat; 38: glycoprotein, AD virus; 39: apolipoprotein A1, rat; 40: glycoprotein, rabies virus; 41: hemagglutinin, human influenza Victoria; 42: hemagglutinin, human influenza Jap; 43: hemagglutinin, avian influenza FPV; 44-50: leukocyte interferon, human; 51: immune interferon, human; 52: fibroblast interferon, human; 53-56: α -immunoglobulin, mouse; 57-59: λ -immunoglobulin, mouse; 60, 61: α -immunoglobulin, mouse; 62: H-chain immunoglobulin, mouse; 63-66: embryonic VH-immunoglobulin, mouse; 67: H-chain immunoglobulin, mouse; 68: trypsinogen 1, canine; 69: trypsinogen 2+3, canine; 70: chymotrypsinogen 2, canine; 71: carboxypeptidase A1, canine; 72: amylase, canine; 73, 74: amylase, mouse; 75: amylase, rat; 76: α -lactalbumin, rabbit; 77: α -lactalbumin, porcine; 78: carboxypeptidase A, rat; 79: ACTH- β -LPH precursor, bovine; 80: ACTH- β -LPH precursor, porcine; 81: ACTH- β -LPH precursor, human; 82: gastrin, porcine; 83: renin, mouse; 84: glycoprotein, trypanosome; 85: somatostatin, catfish; 86-88: somatostatin, anglerfish; 89: calcitonin, rat; 90: glucagon, anglerfish

Protein	Amino acid in position							Reference
	-20	-15	-10	-5	-1*	5	10	
1		M R C F I S L V L G L L A L E V A L A * R N L Q / E H V F N S						[8]
2		M A L H M V L V L S L L P / L L E A * Q N P E / P A N I T L						[9]
3		M D Y Y R K Y A A V I L V M L S M F L H / I L H S * L P D G / D F I I Q G						[10]
4	M A	L S P F L A A V I P L V L L L S R A P / P S A D T * R T T / G H L C G K D						[11]
5		M A A L W L Q S F S L L V L L V V S W P / G S Q A * V A P A / Q H L C G S						[12]
6		M A L W M R L L P L L A L L A L W G P D P A A A * F V N Q H L C G S H						[13]
7		M A L W M R F L P L L A L L V L W E P K P A Q A * F V K Q H L C G P H						[13]
8		M A L W I R F L P L L A L L I L W E P R P A Q A * F V K Q H L C G S H						[14]
9			M K V L I L A C L V A L A / L A * R E Q E E L / N V V G					[15]
10		M R K S I L L V V T I L A L T / L P F L I A * Q E / Q N Q E Q R I C						[15]
11		M M S F V S L L L V G I L / F W A T Q A * E Q / L T K C E V F Q						[15]
12		M K C L L L A L G L A L A C / G V Q A * I I V T / Q T M K G L						[15]
13			M K L L I L T C L V A V A / L A * R P K H P I / K H X G					[15]
14			M K V L M K A C L V A V A / L A * K N T M E H / V S S S					[15]
15			M K C L L Y L A F L F I H V N / C * K F T I V F P / X X X					[16]
16		M Q Y R A L V I A V I L L S T / T V P E V C S * K / S I I D R E R R D						[17]
17		M K F L V N V A L V F M V / V Y I S Y I Y A * A P E P E P A P E P						[18]
18	M N S Q V	S A R K A G T L L L L M M S N L L F / C * Q N V Q T L X / X C X X X C X						[19]
19	M P G S R T S	L L L A F A L L C L P / W L Q E A G A * V / Q T V P L S R L F						[20]
20		M E M F Q G L L L L L L S M G G / T W A * S K E P L / R P R C R						[21]
21		M D Y Y R K Y A A I F L V T L S V F L H / V L H S * A P D V / Q D C P E C						[22]
22		M K L A I T L A L V T L A L L C S P / A S A G * I C P R / F A X V I						[23]
23	M A A D S Q T P W	L L T F S L L C L L W P / Q E A G A * L P A / M P L S S L F						[24]
24	M A T G S R T S	L L L A F G L L C L P / W L Q E G S A * F / P T I P L S R L F						[25]
25	M M A A G P R T S	L L L A F A L L C L P / W T Q V V G A * F / P A M S L S G L F						[26]
26	M M S A K D M V K V M	I V M L A I C / F L A R S D G * K / S V K K R A V S E						[27]
27		M S S R L L L Q L L G F W L F / L S Q P C R A * R / V S E E W M D Q V						[28]
28		M K W V T F L L L L F I S G S / A F S * R G V F R / R E A H K						[29]
29		M K W V T F I S L L L F L F S S / A Y S * R G V F R / R E A H K						[30]
30		M K W V T F L L L L F I S G S / A F S * R G V F R / R E A H K						[31]
31	M R Q A A A P L L P	G V L L L F S I L P A S / Q Q * G G V P G A / I P G G						[32]
32	M A M A G V F V L F S F V L C G / F L P D A A F G * A E V D C S R X X X							[33]
33		M R S L L I L V L C F L P L / A A L G * K V F G / R C E X X X						[33]
34		M K L I L C T V L S L G I / A A V C F A * A P / P K S V I X X X						[33]
35	M P S S V S W G I L L L A G L C C L V / P V S L A * E D P / Q G D A A Q K							[34]
36	M S T V E L S L C L L I M L A V C C Y / E A N A * S Q I C / E L V A H E							[35]
37	M R L S L C L L T I L V V C C Y / E A N G Q T L A * G V C Q A L							[35]
38		M R Y M I L G L L A L A A V / C S A * A K V E / F K E P A						[36]
39		M K A A V L A V A L V F L T G X / X A * X E F X X X / D E P X						[37]
40	M V P Q A L L F V P L L V F P L / C F G * K F P I Y T I L D K							[38]
41		M K T I I A L S Y I F C L V F / A * Q D L P G N D / N N S						[39]
42		M A I I Y L I L L F T A V R / G D Q I C I G * Y / H A N						[39]
43		M N T Q I L V F A L V A V I P / T N A * D K I C L / G H H A V						[39]
44	M A L T F A L L V A L L V L S C / K S S C S V G * C / D L P Q T H S L G							[40]
45	M A L T F Y L M V A L V L S Y / K S S F S * S L G C / D L P Q T H S L G							[40]
46	M A L S F S L L M A V L V L S Y / K S I C S L G * C / D L P Q T H S L G							[40]

THIS PAGE BLANK (USPTO)

Table 1 (continued)

Protein	Amino acid in position							Reference
	-20	-15	-10	-5	-1*	5	10	
47	M A S P F A L M V L V V L S C/K S S C S L G*/C/D L P E T H S L G							[40]
48	M A L S F S L M A V L V L S Y/K S I C S L G*/C/D L P Q T H S L G							[40]
49	M A L P F S L M M A L V V L S C/K S S C S L G*/C/N L S Q T H S L N							[40]
50	M A L P F A L M M A L V V L S C/K S S C S L G*/C/N L S Q T H S L N							[41]
51	M K Y T S Y I L A F Q L C I V L G S/L G*/C Y C Q D P/Y V K E							[42]
52	M T N K C L L Q I A L L L C F S T/T A L S*/M S*/Y N/L L G F L Q							[43]
53	M R A P A Q I F G F L L L L F P G T/R C*/D I Q M T Q/S P S S							[44]
54	M E T D T L L L W V L L L W V P/G S T G*/D I V L/T Q S P A S							[44]
55	M E T D T L L L W V L L L W V P/G S T G*/N I V L/T Q S P A S							[44]
56	M E T D T L L L W V L L L W V P/G A D A*/A P T V/S I F P P S							[44]
57	M A W I S L I L S L L A L S S/G A I S*/Q A V V/T Q E S A L							[44]
58	M A W I S L I L S L L A L S S/G A I S*/Q A V V/T Q E S A L							[44]
59	M A W T S L I L S L L A L C S G/A S S*/Q A V V/T Q E S A L							[44]
60	M G V R M E S H T R V F I F L L L W L S G/T D G*/D I V X X/X X X X X							[44]
61	M D M R A P A Q I F G F L L L L T A I P/G I M S*/D V Q L/Q E S G P G							[44]
62	M K V L S L L Y L L L T A I P/G I M S*/D V Q L/Q E S G P G							[44]
63	M G W S W I F L F L L S G T A G/V H S*/X X X X X/X X X X X							[45]
64	I K W S W I S L F L L S G T A G V/H S*/X X X X X/X X X X X							[45]
65	M E C S W V F L F L L S L T A G/I H C*/X X X X X/X X X X X							[45]
66	M E W S G V F I F L L S V T A G/V Y S*/X X X X X/X X X X X							[45]
67	M G W S F I F L F L L S V T A G V H S*/E V Q L/Q Q S G A E							[46]
68	X X P L L I L A F L X A A V/A*/T P T D D D D/K I V							[47]
69	A L X I I F L A L L X A X/V A*/F P I D D D D/K I V							[47]
70	A F L I L V X A F A L X X V A F X X X V P A I X P V X							[47]
71	X X L I L V F G A L L X A I Y X Q X A F V X X X X X							[47]
72	X X F F L L L X V I G F X/V A*/Q Y A P H X/X X X X							[47]
73	M K F F L L L S L I G F C/W A*/Q Y D P H T/Q Y G R							[47]
74	M K F V L L L S L I G F C/W A*/Q Y D P H T/S D G R							[47]
75	L L S L I G F C Y A*/Q Y D P H T A D G R							[47]
76	M M P L V P L L L V S I V/F P G I Q A*/T Q/L T R C E L T E							[48]
77	M M S F V S L L V V G I L/F P A I Q A*/K Q/F T K C E L S Q							[48]
78	M K R L L I L S L L L E A V/C G*/N E N F V G/H Q V L							[49]

Sequences with unknown cleavage sites

79	M P R L C S S R S G A L L L A L L L Q A S M/E V R G*/W C L E/S S Q C Q D	[50]
80	M A W Q G L L L A A C L L V L P/S T M A*/D C L S/G C S L C A	[51]
81	M A R F L T L C T W L L L L G/P G L L A*/T V R/A E C S Q D C	[52]
82	M Q R L C A Y V L I H V L A L A A C/S E A*/S W K P G/F Q L X X	[53]
83	M D R R R M P L W A L L L W S P C T F S*/L P T G T T F E R I	[54]
84	M V K A I A S L M L L H I W A I E E/I K A*/E R Q A P S V S R T	[55]
85	M S S S P L R L A L A L M C L V S A V G*/V I S*/C G R P H V/V L N S	[56]
86	M V S S S R L R C L L V L L L S L T V S/I S C S F A*/G Q/R D S K L R L L	[57]
87	M K M V S S S R L R C L L V L L L S L T A S/I S C S F A*/G Q/R D S K L R L L	[58]
88	M Q C I R C P A I L A L L A L V L C G P/S V S*/S Q L D R/E Q S D N	[58]
89	M G F L K F S P F L V V S I/L L L Y Q A C G*/L Q A V P L R S T L	[59]
90	M K R I H S L A G I L L V L G L I Q S S/C*/R V L M Q E A/D P S	[60]

indicated by the observed greater proportion of charged residues in this position (Table 2).

A structure of this kind would need some 8 or 9 helical residues added to the 4 or 5 in the extended sheet conformation to reach through the non-polar interior of the membrane, a figure that compares quite well with the length of the uncharged segment of the shortest known signal sequence, 13 residues.

A similar proposal has been made earlier on the basis of an analysis of predicted secondary structures for signal sequences [4] and the model is also in agreement with recent results from an energy-minimization study of the conformation of one particular signal sequence [5].

A signal sequence would thus contain two different and largely independent 'signals': one in the form of a hydrophobic

core, presumably responsible for initiating export and for binding to the 'signal recognition protein' [6], and a second one in the region -5 to -1 conferring processing specificity.

This analysis has been based on a collection of eukaryotic signal sequences and it might be asked how well prokaryotic ones conform to the patterns presented here. Unfortunately, only some 16 prokaryotic signal sequences are known so that too great a reliance on a statistical analysis would hardly be justified at this point. From the limited data available it seems, however, that the only obvious differences are limited to a more pronounced exclusion of charged residues from the prokaryotic sequences (except, of course, at the N terminus), and a strong preference for small, neutral residues in position -6 (data not shown).

THIS PAGE BLANK (USPTO)

Table 2. Number of residues of different physico-chemical character in positions -6 to +2 (cf. Fig. 1 and Table 1)

The cleavage site is located between positions -1 and +1 (aromatic residues): Phe, His, Trp, Tyr; II (charged residues): Asp, Glu, Lys, Arg; III (large polar residues): Asn, Gln; IV (hydrophobic residues): Phe, Ile, Leu, Met, Val; V (small neutral residues): Ala, Cys, Ser, Thr; G: Gly; P: Pro; N.B. Phe is included in groups I and IV since it is both bulky and aromatic (thus excluded from positions -1 and -3) and hydrophobic (thus abundant in the hydrophobic core)

Group	Number of residues in position							
	-6	-5	-4	-3	-2	-1	+1	+2
I	11	6	7	0	21	0	6	7
II	0	6	7	0	9	0	25	15
III	0	6	1	0	13	1	14	11
IV	42	7	19	22	29	0	13	16
V	27	27	22	52	9	57	18	9
G	0	7	21	1	3	18	2	4
P	7	16	6	0	0	0	0	11

Table 3. 'Statistical weights' of the various physico-chemical groups of amino acids in positions -5 to +1, used to calculate the 'processing probability' (see text)

The groups are defined in the legend of Table 2

Position	'Statistical weight' of group						
	I	II	III	IV	V	G	P
+1	1.0	1.0	1.0	1.0	1.0	1.0	0.0
-1	0.0	0.0	0.5	0.0	Ala = 8.0 Ser = 4.0 Cys = 1.0 Thr = 1.0	4.0	0.0
-2	2.0	1.0	1.0	1.0	0.6	1.0	0.0
-3	0.0	0.0	0.0	1.0	3.0	1.0	0.0
-4	1.0	1.0	1.0	1.0	1.0	4.0	1.0
-5	1.0	1.0	1.0	0.7	1.0	1.0	4.0

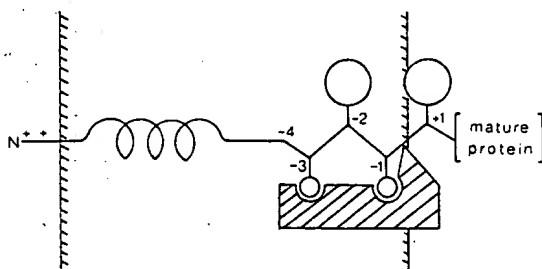


Fig. 1. Proposed signal-sequence-protease complex. The signal sequence spans the membrane as a 'helix + sheet' structure. The small, neutral residues in positions -1 and -3 fit into a pocket in the protease, thereby defining the cleavage site between positions -1 and +1

More generally, prokaryotic signal sequences are richer in Ala and poorer in Leu than eukaryotic ones [7].

In conclusion, it seems that a statistical approach to the problem of signal sequence processing can yield valuable information not readily obtainable from experimental work, and may, with suitable refinements, provide a good basis for schemes predicting the site at which processing of a given pre-protein sequence will take place.

This work was supported by a grant from the Swedish Natural Sciences Research Council.

REFERENCES

1. Strauss, A. W. & Boime, I. (1982) *CRC Crit. Rev. Biochem.* 12, 205-235.
2. von Heijne, G. (1982) *J. Mol. Biol.* 159, 537-541.
3. Chou, P. & Fasman, G. D. (1978) *Annu. Rev. Biochem.* 47, 251-276.
4. Austen, B. M. (1979) *FEBS Lett.* 103, 308-313.
5. Pincus, M. R. & Klausner, R. D. (1982) *Proc. Natl Acad. Sci. USA*, 79, 3413-3417.
6. Walter, P., Ibrahimi, I. & Blobel, G. (1981) *J. Cell Biol.* 91, 545-550.
7. von Heijne, G. (1981) *Eur. J. Biochem.* 116, 419-422.
8. Dandekar, A. M., Robinson, E. A., Appella, E. & Quasba, P. K. (1982) *Proc. Natl Acad. Sci. USA*, 79, 3987-3991.
9. Ricca, G. A. & Taylor, J. M. (1981) *J. Biol. Chem.* 256, 11199-11202.
10. Chin, W. W., Kronenberg, H. M., Dee, P. C., Maloof, F. & Habener, J. F. (1981) *Proc. Natl Acad. Sci. USA*, 78, 5329-5333.
11. Chan, S. J., Emdin, S., Kwok, S. C. M., Kramer, J. M., Falkmer, S. & Steiner, D. F. (1981) *J. Biol. Chem.* 256, 7595-7602.
12. Hobart, P. M., Shen, L.-P., Crawford, R., Pictet, R. L. & Rutter, W. J. (1980) *Science (Wash. DC)* 210, 1360-1363.
13. Bell, G. I., Swain, W. F., Pictet, R., Cordell, B., Goodman, H. M. & Rutter, W. J. (1979) *Nature (Lond.)* 282, 525-527.
14. Lomedico, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R. & Tizard, R. (1979) *Cell*, 18, 545-558.
15. Mercier, J.-C. & Gaye, P. (1980) *Ann. N.Y. Acad. Sci.* 343, 232-251.
16. Lingappa, V. R., Katz, F. N., Lodish, H. F. & Blobel, G. (1978) *J. Biol. Chem.* 253, 8667-8670.
17. Chan, L., Bradley, W. A. & Means, A. R. (1980) *J. Biol. Chem.* 255, 10060-10063.
18. Mollay, C., Vilas, U. & Kreil, G. (1982) *Proc. Natl Acad. Sci. USA*, 79, 2260-2263.
19. McKean, D. J. & Maurer, R. A. (1978) *Biochemistry*, 17, 5215-5219.
20. Sherwood, L. M., Burstein, Y. & Schechter, I. (1979) *Proc. Natl Acad. Sci. USA*, 76, 3819-3823.
21. Birken, S., Fetherston, J., Canfield, R. & Boime, I. (1981) *J. Biol. Chem.* 256, 1816-1823.
22. Fiddes, J. C. & Goodman, H. M. (1979) *Nature (Lond.)* 281, 351-356.
23. Malsky, M. L., Bullock, D. W., Willard, J. J. & Ward, D. N. (1979) *J. Biol. Chem.* 254, 1580-1585.
24. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. (1977) *Nature (Lond.)* 270, 486-494.
25. Martial, J. A., Hallewell, R. A., Baxter, J. D. & Goodman, H. M. (1979) *Science (Wash. DC)* 205, 602-606.
26. Miller, W. L., Martial, J. A. & Baxter, J. D. (1980) *J. Biol. Chem.* 255, 7521-7524.
27. Kronenberg, H. M., McDevitt, B. E., Majzoub, J. A., Nathans, J., Sharp, P. A., Potts, P. A. & Rich, A. (1979) *Proc. Natl Acad. Sci. USA*, 76, 4981-4985.
28. Hudson, P., Haley, J., Cronk, M., Shine, J. & Niall, H. (1981) *Nature (Lond.)* 291, 127-131.
29. Sargent, T. D., Yang, M. & Bonner, J. (1981) *Proc. Natl Acad. Sci. USA*, 78, 243-246.
30. Dugaiczyk, A., Law, S. W. & Dennison, O. E. (1982) *Proc. Natl Acad. Sci. USA*, 79, 71-75.
31. Strauss, A. W., Bennett, C. D., Donohue, A. M., Rodkey, J. A. & Alberts, A. W. (1977) *J. Biol. Chem.* 252, 6846-6855.
32. Karr, S. R. & Foster, J. A. (1981) *J. Biol. Chem.* 256, 5946-5949.
33. Palmiter, R. D., Thibodeau, S. N., Rogers, G. & Boime, I. (1980) *Ann. N.Y. Acad. Sci.* 343, 192-209.
34. Leicht, M., Long, G. L., Chandru, T., Kurachi, K., Kidd, V. J., Mace, M., Davie, E. W. & Woo, S. L. (1982) *Nature (Lond.)* 297, 655-659.
35. Parker, M., Needham, M. & White, R. (1982) *Nature (Lond.)* 298, 92-94.
36. Persson, H., Jörnvall, H. & Zabielski, J. (1980) *Proc. Natl Acad. Sci. USA*, 77, 6349-6353.
37. Stoffel, W., Blobel, G. & Walter, P. (1981) *Eur. J. Biochem.* 120, 519-522.

THIS PAGE BLANK (USPTO)

38. Anilionis, A., Wunner, W. H., & Curtis, P. J. (1981) *Nature (Lond.)* 294, 275-278.
39. Min Jou, W., Verhoeven, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W., Threlfall, G., Barber, C., Carey, N. & Emtage, S. (1980) *Cell*, 19, 683-696.
40. Goeddel, D. V., Leung, D. W., Dull, T. J., Gross, M., Lawn, R. M., McCandliss, R., Seeburg, P. H., Ullrich, A., Yelverton, E. & Gray, P. W. (1981) *Nature (Lond.)* 290, 20-26.
41. Lawn, R., Adelman, J., Dull, T. J., Gross, M., Goeddel, D. & Ullrich, A. (1981) *Science (Wash. DC)* 212, 1159-1162.
42. Gray, P. W. & Goeddel, D. V. (1982) *Nature (Lond.)* 298, 859-863.
43. Taniguchi, T., Mantei, N., Schwarzstein, M., Nagata, S., Muramatsu, M. & Weissman, C. (1980) *Nature (Lond.)* 285, 547-549.
44. Schechter, I., Wolf, O., Kantor, F., Schechter, B. & Burstein, Y. (1980) *Ann. N.Y. Acad. Sci.* 343, 218-231.
45. Givol, D., Zakut, R., Effron, K., Rechavi, G., Ram, D. & Cohen, J. B. (1981) *Nature (Lond.)* 292, 426-430.
46. Sims, J., Rabbitts, T. H., Estess, P., Slaughter, C., Tucker, P. W. & Capra, J. D. (1982) *Science (Wash. DC)* 216, 309-310.
47. Carne, T. & Scheele, G. (1982) *J. Biol. Chem.* 257, 4133-4140.
48. Gaye, P., Hue, D., Raymond, M.-N., Haze, G. & Mercier, J.-C. (1982) *Biochimie (Paris)* 64, 173-184.
49. Quinto, C., Quiroga, M., Swain, W. F., Nikovits, W. C., Standring, D. N., Pictet, R. L., Valenzuela, P. & Rutter, W. J. (1982) *Proc. Natl Acad. Sci. USA*, 79, 31-35.
50. Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang, A. C. Y., Cohen, S. N. & Numa, S. (1979) *Nature (Lond.)* 278, 423-427.
51. Kakidani, H., Furutani, Y., Takahashi, H., Noda, M., Morimoto, Y., Hirose, T., Asai, M., Inayama, S., Nakanishi, S. & Numa, S. (1982) *Nature (Lond.)* 298, 245-249.
52. Comb, M., Seeburg, P. H., Adelman, J., Eiden, L. & Herbert, E. (1982) *Nature (Lond.)* 295, 663-666.
53. Joon Yoo, O., Powell, T. & Agarwal, K. L. (1982) *Proc. Natl Acad. Sci. USA*, 79, 1049-1053.
54. Panthier, J.-J., Foote, S., Chambrault, B., Strosberg, A. D., Corvol, P. & Rougeon, F. (1982) *Nature (Lond.)* 298, 90-92.
55. Rice-Ficht, A. C., Chen, K. K. & Donelson, J. E. (1982) *Nature (Lond.)* 298, 676-679.
56. Magazin, M., Minth, C. D., Funckes, C. L., Deschenes, R., Taviani, M. A. & Dixon, J. E. (1982) *Proc. Natl Acad. Sci. USA*, 79, 5152-5156.
57. Goodman, R. H., Jacobs, J. W., Chin, W., Lund, P. K., Dee, P. C. & Habener, J. F. (1980) *Proc. Natl Acad. Sci. USA*, 77, 5869-5873.
58. Hobart, P., Crawford, R., Shen, L., Pictet, R. & Rutter, W. J. (1980) *Nature (Lond.)* 288, 137-141.
59. Jacobs, J. W., Goodman, R. H., Chin, W. W., Dee, P. C., Habener, J. F., Bell, N. H. & Potts, J. T. (1981) *Science (Wash. DC)* 213, 457-459.
60. Lund, P. K., Goodman, R. H., Dee, P. C. & Habener, J. F. (1982) *Proc. Natl Acad. Sci. USA*, 79, 345-349.

G. von Heijne, Forskningsgruppen för Teoretisk Biofysik, Institutionen för Teoretisk Fysik, Kungliga Tekniska Högskolan, S-100 44 Stockholm. Sweden

THIS PAGE BLANK (USPTO)

The structure of signal peptides from bacterial lipoproteins

Gunnar von Heijne

Department of Molecular Biology, Center for Biotechnology, Karolinska Institutet, Huddinge Hospital-K87, S-141 86 Huddinge, Sweden

Statistical analysis of lipoprotein and non-lipoprotein signal peptides reveals that the two classes differ significantly only in the region close to the signal peptidase cleavage site. This region is apolar and has the consensus sequence LA(G,A)IC in the lipoproteins, but is polar and has small, uncharged residues in positions -3 and -1 in the non-lipoproteins. A simple search for matches to the lipoprotein consensus cleavage site suffices to discriminate between the two groups with close to 100% reliability.

Key words: signal peptide/lipoprotein/consensus sequence/signal peptidase

Introduction

Secretory proteins are normally made with a transiently attached amino-terminal signal peptide that serves as an address label for membrane translocation. Typical signal peptides contain three distinct regions: a positively charged amino-terminal region (n-region), a central, hydrophobic region (h-region) and a more polar, carboxy-terminal region (c-region) that specifies the site of cleavage between the signal peptide and the mature protein (von Heijne, 1985).

In eukaryotic cells, the signal peptide is removed by a multi-subunit enzyme called signal peptidase, located in the endoplasmic reticulum (Mollay, 1985). In bacteria, two distinct signal peptidases have been identified: signal peptidase I (SPase I or leader peptidase, *Lep*; Wickner, 1988) that removes the signal peptide from the great majority of secreted proteins, and signal peptidase II (SPase II, *Lsp*; Innis *et al.*, 1984) that cleaves the signal peptide from a small class of lipoproteins. Cleavage of lipoprotein signal peptides involves, as a first step, the addition of a glyceryl moiety to a Cys residue near the amino terminus, whereupon the signal peptide is cleaved by SPase II on the amino-terminal side of this modified cysteine (Tokunaga *et al.*, 1982).

To assess the differences and similarities between SPase I- and SPase II-cleaved signal peptides, the n-, h- and c-regions from a sample of lipoprotein signal peptides have been compared with a sample of non-lipoprotein signal peptides. It is shown that lipoprotein signal peptides differ in their c-region from non-lipoprotein signal peptides; however, the n- and h-regions of the two classes are indistinguishable in terms of length and amino acid composition.

This observation provides a simple explanation for the results recently reported in this journal by Klein *et al.* (1988), namely that lipoprotein signal peptides tend to be shorter, more hydrophobic and bulky, and have a stronger tendency for turn-conformations in the region immediately downstream of the cleavage site than non-lipoprotein signal peptides. Using multi-dimensional cluster analysis, these authors developed a method whereby some 90% of all lipoprotein signal peptides could be discriminated from some 90% of a control sample of non-lipo-

protein signal peptides. Here, I show that a simple pattern recognition method that searches for an SPase II consensus cleavage site provides a discrimination between lipoprotein and non-lipoprotein signal peptides that is at least as good as the one provided by cluster analysis.

Methods

Thirteen non-homologous lipoprotein signal peptides from Gram-negative bacteria and an additional three from Gram-positive bacteria were extracted from the SIGPEP database (von Heijne, 1987; Table I). Although the site of SPase II cleavage has not been determined in all cases, the 'SPase II consensus' sequence (see below) allowed unambiguous assignments of this site for all sequences. Twenty-eight non-homologous non-lipoprotein signal peptides with known SPase I cleavage sites from *Escherichia coli* were analysed in parallel (these sequences, as well as the SIGPEP database, may be obtained from the author). For the analysis of the c-regions, the signal peptides were aligned from their site of cleavage (between positions -1 and +1), and the position-specific amino acid counts were obtained. The border between the n- and h-regions was defined to be located immediately downstream of the most C-terminal charged residue (Arg, Lys, Asp, Glu) of the n-region; the n- and h-regions thus defined were analysed in terms of length and amino acid composition. Two-sided *t*-test and χ^2 analysis were used to assess the statistical significance of differences between the two groups.

Results

Since signal peptides from Gram-positive and Gram-negative bacteria are different in terms of the lengths and amino acid compositions of the n-, h- and c-regions (von Heijne and Abrahmsén, 1989), lipoprotein signal peptides from Gram-negative bacteria were first compared with non-lipoprotein signal peptides from *E. coli*.

As shown by Klein *et al.* (1988) there is a statistically significant difference ($P < 10^{-3}$) in the mean signal peptide length between these two classes, with the lipoprotein sequences being on average 5 residues shorter (19.2 versus 24.1 residues; Figure 1). To find out which of the three 'typical' regions contribute to this difference, plots of the mean frequency of hydrophobic (ACFILMVW) and turn-promoting (DGNPS) (Levitt, 1978) residues as a function of the position relative to the cleavage site were made (Figure 2). In addition, the mean lengths and overall amino acid compositions of the n- and h-regions in the two samples were calculated.

In the non-lipoprotein signal peptides, the average c-region is 6 residues long and extends from residue -6 to residue -1 (Figure 2), whereas the average lipoprotein c-region cannot be distinguished from the h-region on the basis of hydrophobicity alone (although lipoprotein signal peptides seem to have a weak tendency to have turn-promoting residues in positions -7 to -5). However, all lipoproteins have a SPase II 'consensus' sequence LA(G,A)IC (see below) in positions -3 to +1. The mean

THIS PAGE BLANK (USPTO)

Table I. Lipoprotein signal peptides.

Protein	Sequence	Ref.
<i>Escherichia coli</i> Lpp	MKATKLVLGAVILGSTLLAGACSSNAKIDQS	1
<i>Escherichia coli</i> NlpA	MKLTTHHLRTGAALLLAGAGCDQSSSDAKH	2
<i>Escherichia coli</i> TraT	MKMKKLLMMVALMSSTLALSGACGAMSTAIKK	3
<i>Escherichia coli</i> Pal	MQLNKVLKGLMIALPVMIAAACSNNKNASND	4
<i>Escherichia coli</i> ColE3 lysis protein	MKKITGIIILLLLAVIILSAACQANYIRDVQ	5
<i>Escherichia coli</i> RplA	MRKQWLGCICIAAGMLAAACTSDDGQQQT	6
<i>Escherichia coli</i> RplB	MRYLATLLLSLAVLITAGACGWHLRDITQ	6
<i>Escherichia coli</i> Col E1 lysis protein	MRKRFFVVGIFAINLLVGACQANYIRDVQ	7
<i>Haemophilus</i> Pcp	MKKTNMALALLVAFSVTGACANTDIFSGD	8
<i>Haemophilus</i> Pal	MNKFVKSLLVAGSVAALAAACSSSNNDAAAG	8
<i>Rickettsia</i> surface antigen	MKLLSKIMIIALATSMQLQAACNGPGGMNKQ	9
<i>Rhodospseudomonas viridis</i> cytochrome	MKQLIVNSVATVALASLVAGACFEPPTATT	10
<i>Klebsiella pullulanase</i>	MLRYTCNALFLGSLILLSGACDSSSSSSSS	11
<i>Staphylococcus aureus</i> Bla	MKKLI FLIVIALVL SAACNSNHAKELN	12
<i>Bacillus licheniformis</i> Bla	MKLWFS TLK LKKA AAVLL FSCVALAGACANNQTNASQ	13
<i>Bacillus cereus</i> BlaZ	MFVLNKFFTNSHYKKIVPVVLLSCATLIGACSNSTQSES	14

The SPase II cleavage site is indicated by A. References: 1, Nakamura *et al.* (1980); 2, Yu *et al.* (1986); 3, Ogata *et al.* (1982); 4, Chen and Henning (1987); 5, Watson *et al.* (1984); 6, Takase *et al.* (1987); 7, Waleh and Johnson (1985); 8, Deich *et al.* (1988); 9, Anderson *et al.* (1987); 10, Weyer *et al.* (1987); 11, Chapon and Raibaud (1985); 12, McLaughlin *et al.* (1981); 13, Nielsen and Lampen (1982); 14, Hussain *et al.* (1987).

lengths of the h + c regions of the two samples are 14.7 (lipoproteins) and 18.1 (non-lipoproteins) residues, again a statistically significant difference ($P < 10^{-3}$). However, if the average c-region lengths (3 and 6 residues) are subtracted from the h + c region lengths, the resulting mean h-region lengths are 11.7 and 12.1 residues; these mean lengths are not significantly different ($P > 0.1$). The overall amino acid compositions of the h-regions are identical to within statistical error (Figure 3).

The mean n-region length is 4.5 for the lipoprotein signal peptide and 6.0 for the non-lipoprotein *E. coli* sample. The mean n-region net charge is +1.9 and +2.0 respectively. These differences are not statistically significant ($P > 0.1$). The overall amino acid compositions of the n-regions also do not differ significantly (Figure 3).

An obvious difference between the two samples, already noted by Klein *et al.* (1988), is the high incidence of turn-promoting residues immediately downstream of the cleavage site (positions +2 to +6) in the lipoprotein sample. In contrast, among the non-lipoprotein signal peptides, turn-promoting residues tend to be found around positions -6 to -4 and in position +2 (Figure 2; von Heijne, 1983).

Only three lipoprotein signal peptides from Gram-positive bacteria are known (Table I). They seem to have longer and more highly charged n-regions than lipoprotein signal peptides from Gram-negative bacteria, but this is observed also for non-lipoprotein signal peptides (von Heijne and Abrahmsén, 1989). Again, the only obvious difference between SPase I- and SPase II-cleaved signal peptides is the design of the c-region.

Discussion

From the results presented above, it is apparent that lipoprotein and non-lipoprotein signal peptides differ in their c-regions, but not in their n- and h-regions. One would thus expect them to follow basically the same route through the secretory pathway, up to the point of signal peptide cleavage. This indeed seems to be the case, since lipoproteins are sensitive to most of the mutations (Wu and Tokunaga, 1986), and since when mutations

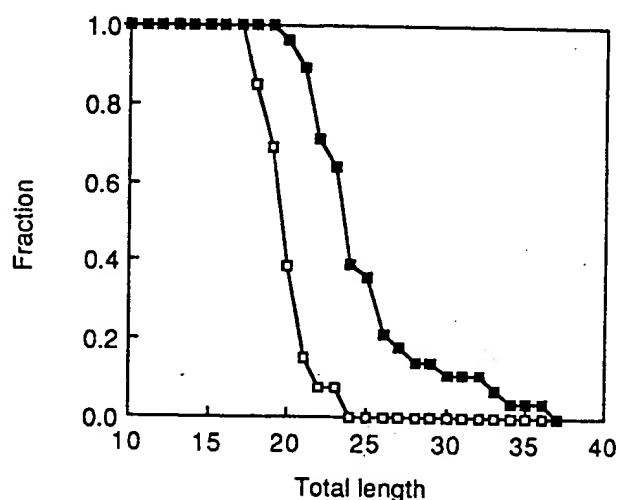


Fig. 1. Cumulative length-distribution of signal peptides from lipoproteins (—□—) and non-lipoproteins (—■—).

that prevent cleavage by SPase II are introduced into the lipoprotein c-region, 'cryptic' cleavage sites most likely recognized by SPase I are often uncovered (Mézes *et al.*, 1983; Hayashi *et al.*, 1984, 1986; Ghayeb *et al.*, 1985). In all such cases studied to date, the cryptic sites are located downstream of the SPase II-site. A similar situation has been observed with lipoprotein mutants expressed in yeast (Pines *et al.*, 1988). This is consistent with the fact that SPase I c-regions typically contain a stretch of rather polar amino acids, whereas SPase II c-regions are composed largely of apolar residues that SPase I may view as being part of the h-region.

The simple observation that the c-region of lipoprotein signal peptides is apolar and somewhat shorter than the c-region of SPase I-cleaved signal peptides explains most of the differences between the two classes of peptides found by Klein *et al.* (1988). Thus, the greater mean hydrophobicity (taken over the whole signal peptide), the greater mean side-chain volume, the smaller overall Ser + Thr + Pro content, and the shorter mean length

THIS PAGE BLANK (USPTO)

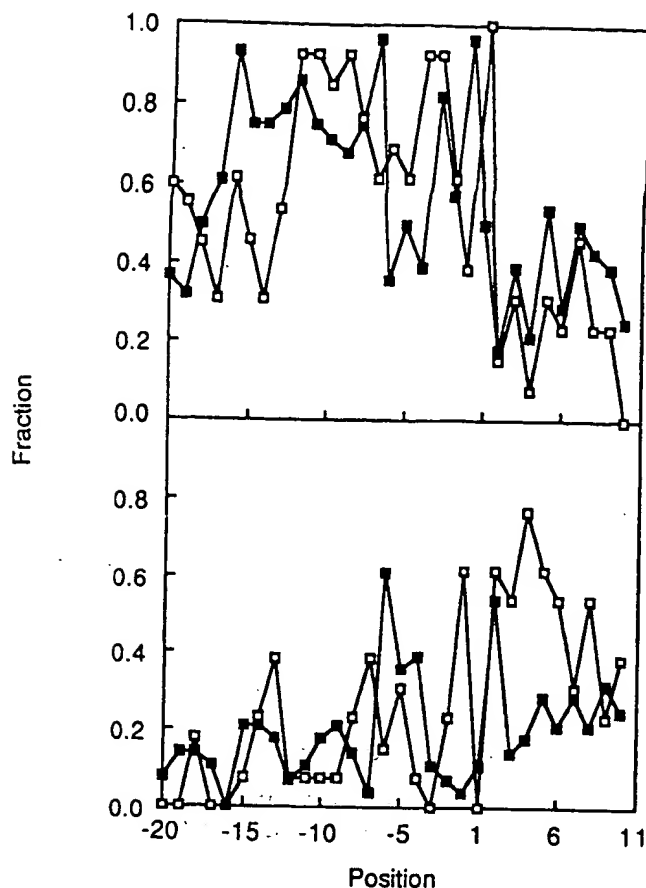


Fig. 2. Incidence of apolar (ACFILMVW, top panel) and turn-promoting (DGNPS, bottom panel) residues as a function of position relative to the cleavage site for lipoprotein (—□—) and non-lipoprotein (—■—) signal peptides.

of the former group are immediate consequences of this fact. The turn-propensity, both by Chou–Fasman (1978) and mean energy (Oobatake and Ooi, 1977) analysis, was found to peak in the region downstream of the SPase II cleavage sites but around position -5 relative to the SPase I sites, again reflecting the difference in the c-region design.

This suggests that the best discrimination between the two classes of signal peptides should be obtained by focusing on the c-region proper rather than on averages calculated over the whole peptide. From the available data on wild-type and mutant lipoprotein signal peptides, it seems that the Cys in position $+1$ is absolutely required (Inouye *et al.*, 1983), that Ala, Gly and Ser are acceptable in position -1 (Pollitt *et al.*, 1986; although Ser $_{-1}$ has never been found in a wild-type sequence), and that Ala and Leu are strongly preferred but not absolutely required in positions -2 and -3 respectively. A simple 'consensus'-type search where putative SPase II cleavage sites are identified by scanning, say, residues 10–20 counting from the most C-terminal of the charged n-region residues (the precise ends of the search-region are not important for the results; in known lipoprotein signal peptides the distance from the n-region to the cleavage site is between 13 and 18 residues) with the four-residue consensus (L,V,D)(A,S,T,G)(G,A)C and requiring at least one match to the first two positions and precise matches to the final two, suffices to discriminate between *all* of the known SPase I- and SPase II-cleaved signal peptides in the entire SIGPEP database (a total of some 200 entries, including homologous lipoprotein sequences not shown in Table I).

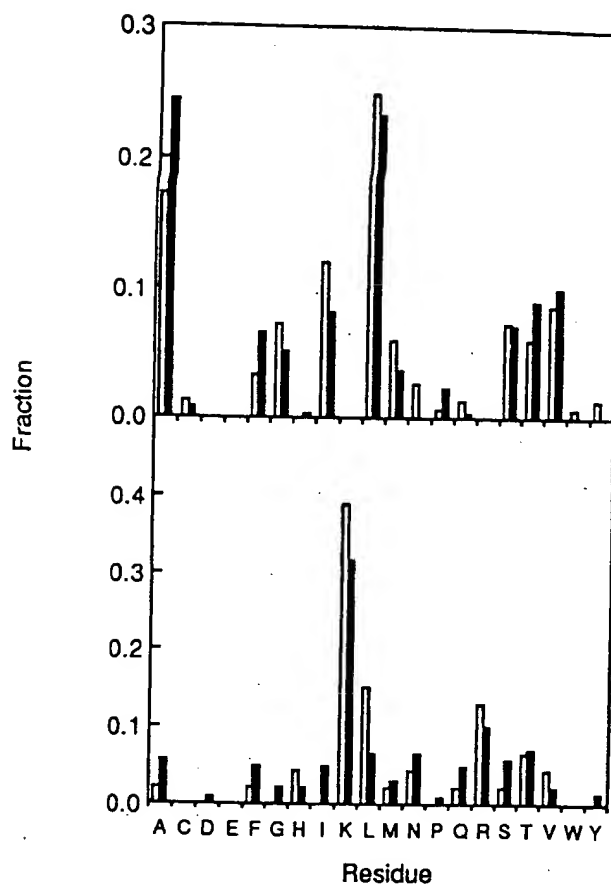


Fig. 3. Overall amino acid composition of h-regions (top panel) and n-regions (bottom panel) from lipoproteins (white bars) and non-lipoproteins (black bars).

Only two signal peptides in the SIGPEP collection matching the SPase II consensus are not known to be lipoproteins: the photosystem II protein WoxA from the cyanobacterium *Anacystis nidulans* (Kuwabara *et al.*, 1987) and the XP55 protein from *Streptomyces lividans* (Burnett *et al.*, 1987). The signal peptidase cleavage site has not been determined for the WoxA protein, and it has sites that are compatible both with SPase I and SPase II cleavage: ...LTA|CSSGPTAA|DL... The secreted XP55 protein does not have a well-defined amino terminus and could possibly result from late proteolytic processing of an initially glycerol-bound form as is the case for, for example, *Bacillus licheniformis* penicillinase (Mézes *et al.*, 1983); the putative SPase II site in XP55 reads ...ATA|CSAPT...

Finally, the greater hydrophobicity of the SPase II c-region, which effectively forms a continuation of the h-region but with a specific consensus sequence, is a hint that the active site of this enzyme may be embedded in the lipid bilayer. The SPase I active site, on the other hand, would by the same token be expected to be located just outside of the bilayer.

Acknowledgements

This work was supported by a grant from the Swedish Natural Sciences Research Council.

References

- Anderson, B.E., Régnery, R.L., Carlone, G.M., Tzinabos, T., McDade, J.E., Fu, Z.Y. and Bellini, W.J. (1987) *J. Bacteriol.* **169**, 2385–2390.
- Burnett, W.V., Henner, J. and Eckhardt, T. (1987) *Nucleic Acids Res.* **15**, 3926.

THIS PAGE BLANK (USPTO)

- Chapon,C. and Raibaud,O. (1985) *J. Bacteriol.*, 164, 639-645.
- Chen,R. and Henning,U. (1987) *Eur. J. Biochem.*, 163, 73-77.
- Chou,P. and Fasman,G.D. (1978) *Adv. Enzymol.*, 47, 45-148.
- Deich,R.A., Metcalf,B.J., Finn,C.W., Farley,J.E. and Green,B.A. (1988) *J. Bacteriol.*, 170, 489-498.
- Ghrayeb,J., Lunn,C.A., Inouye,S. and Inouye,M. (1985) *J. Biol. Chem.*, 260, 10961-10965.
- Hayashi,S., Chang,S.-Y. and Wu,H.C. (1984) *J. Biol. Chem.*, 259, 10448-10454.
- Hayashi,S., Chang,S.-Y., Chang,S. and Wu,H.C. (1986) *J. Bacteriol.*, 165, 678-681.
- Hussain,M., Pastor,F.I.J. and Lampen,J.O. (1987) *J. Bacteriol.*, 169, 579-586.
- Innis,M.A., Tokunaga,M., Williams,M.E., Loranger,J.M., Chang,S.-Y., Chang,S. and Wu,H.C. (1984) *Proc. Natl. Acad. Sci. USA*, 81, 3708-3712.
- Inouye,S., Franceschini,T., Sato,M., Itakura,K. and Inouye,M. (1983) *EMBO J.*, 2, 87-91.
- Klein,P., Somorajai,R.L. and Lau,P.C.K. (1988) *Prot. Engng.*, 2, 15-20.
- Kuwabara,T., Reddy,K.J. and Sherman,L.A. (1987) *Proc. Natl. Acad. Sci. USA*, 84, 8230-8234.
- Levitt,M. (1978) *Biochemistry*, 17, 4277-4285.
- McLaughlin,J., Murray,C.L. and Rabinowitz,J. (1981) *J. Biol. Chem.*, 256, 11283-11291.
- Mézes,P.S.F., Wang,W., Yeh,E.C.H. and Lampen,J.O. (1983) *J. Biol. Chem.*, 258, 11211-11218.
- Mollay,C. (1985) *The Enzymology of Post-translational Modification of Proteins*. Academic Press, New York, Vol. 2, pp. 1-23.
- Nakamura,K., Pirtle,R.M., Takeishi,K. and Inouye,M. (1980) *J. Biol. Chem.*, 255, 210-216.
- Nielsen,J.B.K. and Lampen,J.O. (1982) *J. Biol. Chem.*, 257, 4490-4495.
- Ogata,R.T., Winters,C. and Levine,R.P. (1982) *J. Bacteriol.*, 151, 819-827.
- Oobotake,M. and Ooi,T. (1977) *J. Theor. Biol.*, 67, 567-584.
- Pines,O., Lunn,C.A. and Inouye,M. (1988) *Mol. Microbiol.*, 2, 209-217.
- Pollitt,S., Inouye,S. and Inouye,M. (1986) *J. Biol. Chem.*, 261, 1835-1837.
- Takase,I., Ishino,F., Wachi,M., Kamata,H., Doi,M., Asoh,S., Matsuzawa,H., Ohta,T. and Matsuhashi,M. (1987) *J. Bacteriol.*, 169, 5692-5699.
- Tokunaga,M., Tokunaga,H. and Wu,H.C. (1982) *Proc. Natl. Acad. Sci. USA*, 79, 2255-2259.
- von Heijne,G. (1983) *Eur. J. Biochem.*, 133, 17-21.
- von Heijne,G. (1985) *J. Mol. Biol.*, 184, 99-105.
- von Heijne,G. (1987) *Prot. Seq. Data Anal.*, 1, 41-42.
- von Heijne,G. and Abrahmsén,L. (1989) *FEBS Lett.*, 244, 439-446.
- Waleh,N.S. and Johnson,P.H. (1985) *Proc. Natl. Acad. Sci. USA*, 82, 8389-8393.
- Watson,R.J., Lau,P.C.K., Vernet,T. and Visentin,L.P. (1984) *Gene*, 29, 175-184.
- Weyer,K.A., Lottspeich,F., Gruenberg,H., Lang,F., Oesterhelt,D. and Michel,H. (1987) *EMBO J.*, 6, 2197-2202.
- Wickner,W. (1988) *Biochemistry*, 27, 1081-1086.
- Wu,H.C. and Tokunaga,M. (1986) *Curr. Top. Microbiol. Immunol.*, 125, 127-157.
- Yu,F., Inouye,S. and Inouye,M. (1986) *J. Biol. Chem.*, 261, 2284-2288.

THIS PAGE BLANK (USPTO)

Species-specific variation in signal peptide design

Implications for protein secretion in foreign hosts

Gunnar von Heijne* and Lars Abrahmsén

*Department of Molecular Biology, Center for Biotechnology, Karolinska Institutet, Huddinge Hospital-K87, S-141 86 Huddinge and Department of Biochemistry, Royal Institute of Technology, S-100 44 Stockholm, Sweden

Received 3 January 1989

Secretory signal peptides from individual prokaryotic and eukaryotic species have been analyzed, and the lengths and amino acid compositions of the positively charged amino-terminal region, the central hydrophobic region, and the carboxy-terminal cleavage-region have been compared. We find distinct differences between species in all three regions. Implications for protein secretion in foreign hosts are discussed.

Protein secretion; Signal peptide; Leader peptide

1. INTRODUCTION

The mechanism of protein secretion is highly conserved throughout the living world. In the great majority of cases, secretory proteins are made with an amino-terminal extension, a signal peptide, that targets the precursor to translocation sites on the appropriate membrane. During or shortly after translocation, the signal peptide is removed by a signal peptidase.

Signal peptides from bacteria to plants and mammals share a common design [1]. The 'canonical' signal peptide is characterized by a short, positively charged amino-terminal region (n-region) followed by a central hydrophobic region (h-region) and a more polar carboxy-terminal region that contains the cleavage site (c-region). The signal peptidase apparently recognizes a '(-3, -1)-pattern' with small, uncharged residues in positions -3 and -1 relative to the cleavage site [2].

Correspondence address: G. von Heijne, Department of Molecular Biology, Center for Biotechnology, Karolinska Institutet, Huddinge Hospital-K87, S-141 86 Huddinge, Sweden

Despite these conserved features, previous comparisons have shown that eukaryotic signal peptides differ in detail from prokaryotic ones [3], and that signal peptides from Gram-positive bacteria tend to be longer than those from Gram negative species [4]; also, signal peptides from one organism do not always function efficiently when expressed in foreign hosts [5,6]. With the much larger database now available, we have undertaken a thorough comparative study of signal peptides from different organisms in an attempt to define more precisely the species-specific variations in signal peptide design.

2. METHODS

All signal peptides were selected from the current version of the SIGPEP database [7], which holds a total of about 200 prokaryotic and 900 eukaryotic sequences. Analysis of the h- and c-regions was carried out for signal peptides with known cleavage sites from *E. coli* (28 sequences), *Bacillus* (16 sequences), *Staphylococcus* (6 sequences), *Streptomyces* (6 sequences), *Homo sapiens* (147 sequences), and a collection of plant signal peptides (22 sequences). In the analysis of the n-regions, additional sequences where the signal peptidase cleavage site is not known were included (31 from *E. coli*, 13

THIS PAGE BLANK (USPTO)

Table 1

Bacterial and yeast signal peptides analyzed in this paper

Signal peptide	Ref.
<i>Bacillus</i>	
α -Amylase	MFAKRFTSLPLFAGFLLFYLVLAGPAAASA:ETANKSNELT 20
α -Amylase	MIQKRKRTVSFRLVLMCTLLFVSLPITKTA:VNGTLMQFEW 21
β -Lactamase II	MKKNTLLKVGLCVGLLGTIQFVSTISSVQA:SQKVEKTVIK 22
Middle wall protein	MKKVVNSVLASALALTVPMAFA:AEAAATTTAP 23
Extracellular amylase	MKMRTGKGFLSILLAFLLVITSIPFTLVDEA:HHNGTNGTMM 24
α -Amylase	MKQHKRLYARLLPLLALIFLLPHSAAAA:ANLNGTLMQY 25
Cyclodextrin glucanotransferase	MKRFMKLTAVWTLWLSLTGLLSPVHA:APDTSVSNKQ 26
Cyclodextrin glucanotransferase	MKSRYKRLTSLALSLSMALGISLPAWA:SPDTSVDNKV 27
Alkaline cellulase	MLRKKTKQLISSILVLLLSLFPALAA:EGNTREDNFK 28
α -Amylase	MLTFHRIIRKGMFLLAFLLTALLFCPTGQPAKA:AAFNGTMMQ 29
Levansucrase	MNIKKFAKQATVLTFTTALLAGGATQAF:KETNQKPYKE 30
Outer wall protein	MNKKVLSVLSSTLVASVAASAFA:APKDGIIYIGG 31
Xylanase	MNLRLRLLLFVMCIGLTLITAVPAHA:RTITNNEMGN 32
Subtilisin E	MRSKKLWISLLFALTIFTMAFSNMSAQA:AGKSSTEKKY 33,34
β -Amylase	MTLYRSLWKKGCMILLSLVLSLTAFIGSPNTASA:AVADDFQASV 35
Sphingomyelinase	MKGKLLKGVLSLGVGLGALYSCTSAQA:EASTNQNDTL 36
β -Lactamase II	MFVLNKFNTSHYKKIVPVLLSCATLIGCSNSNTQSES 37
Neutral protease	MGLGKKLSSAVAASFMSTLISLPGVQAAENPQLKENL 38
Amylase	MKGKKWTALALTPLAASLTGVDATVHKGKAPT 39
Levanase	MKKKVLALAAAITVAPLQSVAFAHENDGGSKIK 40
β -Lactamase	MKKRLIQVMIMFTLLLTMAFSADAADSSYYDEDY 41
β -Lactamase	MKLWFTSLKLLKAAAVLLFSCVALAGCANNQTNASQ 41
β -Lactamase I	MKNKKMLKIGMVCVGLGLSITSLVFTTGGALQVEAKEKTG 42
β -Lactamase II	MKNTLLKLGVCVSLGITPVSTISSVQAERTVEHKVVK 43
Endo- β -1,4-glucanase	MKRSIIFITCLLITLLTMGGMIASPASAAAGTKTPVAKN 44
Ribonuclease	MMKMEGIALKKRLSWISVCLLVLSAAGMLFSTAATKTESHKA 45
Subtilisin Carlsberg	MMRKKFWLGLMTAFMLVFTMAFSDSASAAQPAKNVEKD 46
Neutral protease	MNKRAMLGAIGLAFGLLAAPIGASAKGESIVWNEQ 47
β -Glucanase	MPYLKRVLLLVLTGLFMSLFAVTATASAKTGGSFDDPF 48
<i>Staphylococcus</i>	
Nuclease	MAISNVSKGQYAKRFFFFATSCLVLTLLVVSSLSANA:SQTDNGVNRS 49
Exfoliative toxin B	MDKNMFKKIIAASIFTISLPVIFESTLQA:KEYSAEEIRK 50
Protein A	MKKKNIYSIRKLGVGIASVTLGLLISGVTTPAANA:AQHDEAQNA 51
Enterotoxin A	MKKTATLTLFIALTLTTSPLVNG:SEKSEEINEK 52
α -Toxin	MKTRIVSSVTTLLLSILNPNVAGA:ADSDINIKTG 53
Enterotoxin C1	MNKSRLFSCVILIFALILVLFPTNVA:ESQPDPTPDE 54
Lipase	MKETKHQHTFSIRKSA'GAASFMVASCIFVIGGVVAEANDSTTQT 55
Serine protease	MKGKFLKVSSLFVATLTATLVSSPAANALSSKAMDHP 56
Lysostaphin	MKKTNNYYTRPLAIGLSTFALASIVYGGIQNE 57
Staphylokinase	MLKRSLLFLTLLLSLFSSTITNEVSASSSFDKGGKYK 58
Lipase	MLRGQEERKYSIRKYSIGVSVLAATMFV'VSSHEAQASEKTSTNAAA 59
Exfoliative toxin A	MNSKIISKVLLSLSLFTVGASAFVIQDELMEK 60
<i>Streptococcus</i>	
Scarlet fever toxin	MENNKEVLKKMVFFVLMKFLGLTILPKGIC:STRPKPSQLQ 61
Protein G	MEKEKKVKYFLRKSAFGLASVSAFLVGSTVFA:VDSPIEDTPI62,63
Type 6 M protein	MAKNNTNRHYSRLKKGTSVAVALSVIGAGLVNNTNEVSA:RVFPRGTVEN 64
Streptokinase	MKNYLSFGMFALLFALTFTVNSVQA:IAGPEWLLDR 65
Fructosyltransferase	METKVRKKMYKKGFVWVATITTAMLTIGLSSVQADEANSTQVSS 66
Streptolysin O	MSNKKTFKKYSRVAGLLTAALHGNLVTANAESNKQNTAST 67
<i>Streptomyces</i>	
α -Amylase	MARRLATASLAVLAAAATALTAPTPAAA:APPGAKDVT 68

THIS PAGE BLANK (USPTO)

Table 1 (continued)

Signal peptide	Ref.
SapA	MKRSMQAVGATLTAVGAIGAGLLVTAPAAGAATAGAT:ASYNGVCGSG 69
Streptavidin	MRKIVVAAIAVSLTTVSITASASA:DPSKDSKAQV 70
Agarase	MVNRRDLIKWSAVALGAGAGLAGPAPAAHA:ADLEWEQYPV 71
DD-peptidase	MVSGTVGRGTALGAVLLALLAVPAQAGTAAA:ADLPAPDDTG 72
α -Amylase	MQQRSVLGGTLGIVAAAAATVAPWPSQA:TTPGQKTVTA 73
Cellulase	MENPRTTPTPLRRRSERRARGGRVLTALTGVTLTAGLAIAPATGASPSAPPASP 74
EndoH	MFTPVRRRVRTAALALSAALVVGSTAAASGASATPSPAPA 75
β -Lactamase*	MHPSTSRPSRRTLLTATAGAALAAATLVPGTAHASSGGRGHGSGVSDAERRLAGL 76
β -Galactosidase	MPHSPVSPAESPAPQGRPRPVVSRRLLEGGAAVLGALALSASPLTAQAAVRRAADEPPEWDF 77
Protease A	MTFKRFSPLSSTRYARLLAVASGLVAAAALATPSAVAPEAESKATV 78
Protease B	MRIKRTSNRSNAARRVRTTAVLAGLAAVAALAVPTANAETPRTFSANQ 78
XP55	MTARRTRWTRRTDRSLPIRSAAAFAAGATACAPTGGGGDGGTEAAE 79
Yeast	
α -Galactosidase	MFAFYFLTACISLKGVFV:VSPSYNGLGL 80
Acid phosphatase	MFKSVVYSILAASLANA:GTIPLGKLAD 81
Carboxypeptidase Y	MKAFTSLCGLGLSTTLAKA:ISLQRPL... 82
28 kDa killertoxin	MKIYHIFSVCYLITLCA:AATTAREEFF 83
Invertase	MLLQAFLLLAGFAAKISA:SMNETSDRP 84
Mating factor α -1	MRFPSIFTAVLFAASSALA:APVNTTTEDE85,86
PEP4	MFSLKALLPLALLVSANQVAAKVHKAKIYKH 87
Mating factor α -2	MKFISTFLTILAAVSVTASDEDIAQVPA 88
BARI	MSAINHLCLKLILASFAIINTITALNDGTGHLE 89
K1 toxin	MTKPTQVLVRSVSILFFITLLHLVVALNDVAGPAET 90
Glucoamylase	MVGLKNPYHTMQRPFLAYLVLSLLFNSALGFPTALVPRGS 91
Killer plasmid ORF2	MNIFYIFLLSFVQGLEHTRRGSL 92

Known signal peptidase I cleavage sites are indicated by :

from *Bacillus*, 6 from *Staphylococcus*, 6 from *Streptococcus*, 7 from *Streptomyces*, 12 from yeast, and 9 from plants). The sequences from Gram-positive bacteria and from yeast are listed in table 1. Listings of the other samples can be obtained from G.v.H., who also distributes the SIGPEP database in Macintosh format.

Statistical significance was assessed by χ^2 and two-sided *t*-test analysis.

3. RESULTS

Cumulative distributions of the overall lengths of the different signal peptide samples are presented in fig. 1. There is a clear gradation from the short *Homo* and plant sequences (mean length = 22.5 and 23.9 residues) and the *E. coli* sequences (mean length = 24.1) to the much longer signal peptides from Gram-positive bacteria (mean length = 29–31). The differences in mean length between these two groups are statistically significant ($p < 0.005$ by two-sided *t*-test).

3.1. The c-region

Fig. 2 shows plots of the incidence of

hydrophobic (ACFILMV) and turn-promoting (DGNPS) [8] residues when the different samples are aligned with coincident cleavage sites. The mean c-region length, as read off from the figure, is five residues for the *Homo* and plant samples, six residues for the *E. coli*, and eight residues for the *Bacillus* samples. For *Staphylococcus*, *Streptococcus*, and *Streptomyces*, the number of signal peptides with known cleavage sites is too small to make reliable estimates of the c-region length.

As for the cleavage site, the $(-3, -1)$ -rule is faithfully obeyed in all samples (not shown).

3.2. The h-region

We define the border between the n- and h-regions as being located immediately after the most C-terminal charged residue in the n-region. In some cases, this residue is followed by one or more uncharged but polar residues before the first strongly hydrophobic residue is encountered, but for simplicity we adhere to the above definition in these cases as well. In practice, this will not matter

THIS PAGE BLANK (USPTO)

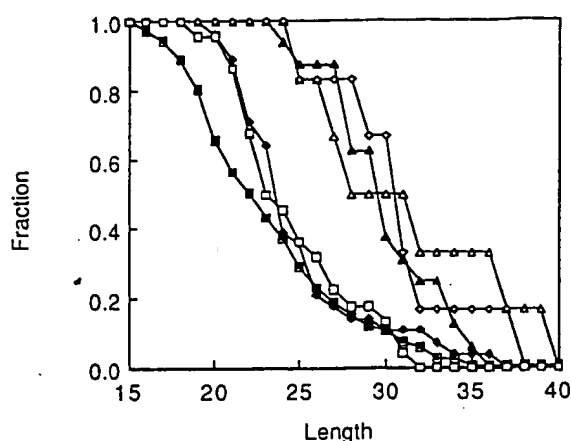


Fig. 1. Cumulative distributions of the overall lengths of signal peptides with known cleavage sites from different species [*Homo* (□), plants (■), *Bacillus* (▲), *Staphylococcus* (Δ), *Streptomyces* (○), *E. coli* (♦)].

for our qualitative conclusions and will have only a minor influence on the mean lengths and amino acid compositions calculated below.

For the combined h- and c-regions, we find that the total mean lengths are 17.6 for *Homo*, 20.0 for the plant sample, 18.2 for *E. coli*, 21.8 for *Bacillus*, 23.0 for *Staphylococcus*, and 24.5 for *Streptomyces*. Subtracting the c-region lengths found above, we thus calculate mean h-region lengths ranging from around 12 residues (*Homo* and *E. coli*) to 15 or more residues for the Gram-positive signal peptides. In terms of overall amino acid composition, the eukaryotic h-regions are relatively rich in Leu (40%) and contain less Ala (10%); h-regions from *E. coli* and all the Gram-positive bacteria contain 25–35% Leu and 10% (*Bacillus*, *Staphylococcus*) or 30% (*E. coli*, *Streptomyces*) Ala.

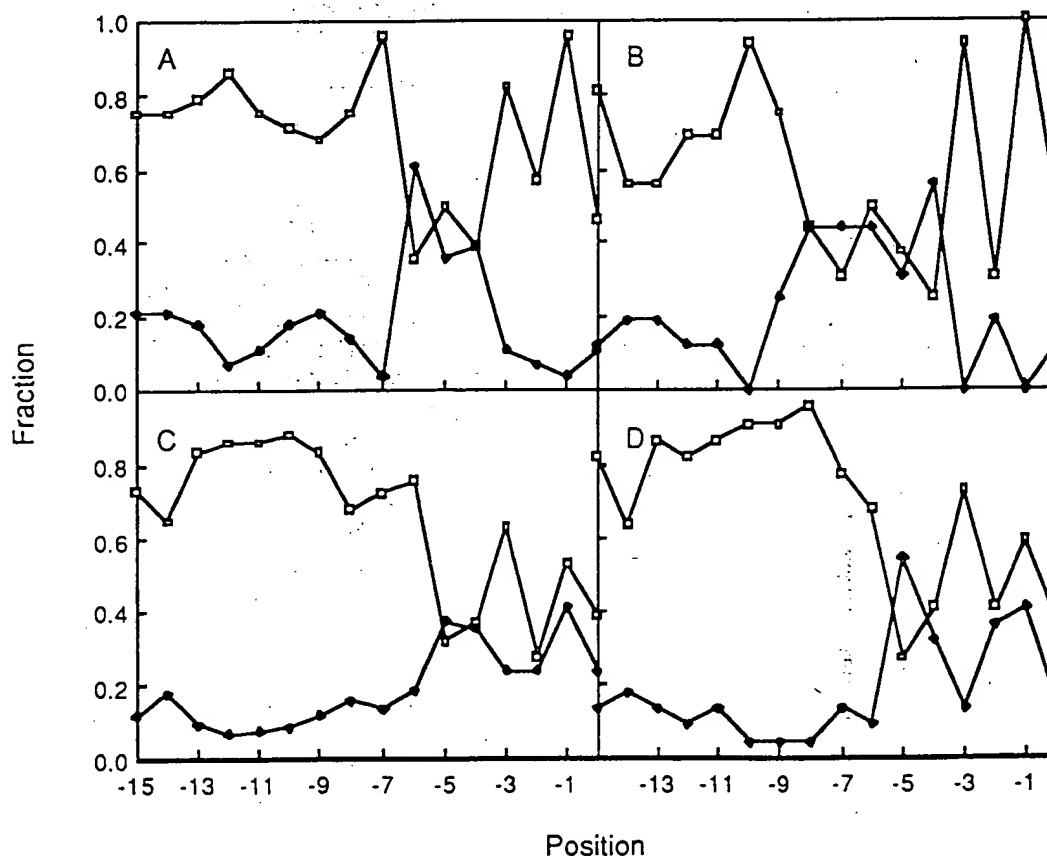


Fig. 2. Distribution of hydrophobic [Ala + Cys + Phe + Ile + Leu + Met + Val (□)] and turn-promoting [Asp + Gly + Asn + Pro + Ser (■)] residues relative to the signal peptidase I cleavage site (between position -1 and +1) for different species [(A) *E. coli*, (B) *Bacillus*, (C) *Homo*, (D) plants].

THIS PAGE BLANK (USPTO)

3.3. The n-region

The length and net charge distributions for the n-region follow similar trends, with the *Homo*, yeast, and plant samples being shortest (mean length 4–5 residues) and carrying the smallest net charge (mean net charge around +0.8, not counting the amino group on the initiator Met), the *E. coli* sample being intermediate (5.5 residues, +2.0), and the *Bacillus* and *Staphylococcus* samples being longer (7–8 residues) and more highly charged (+3.0 and +2.8). The n-regions from *Streptococcus* and *Streptomyces* are particularly long (12 residues) and have mean net charges of +4.3 and +3.5, respectively.

In terms of overall amino acid composition, the n-regions from *Streptomyces* stand out from the other samples by virtue of their very high Arg (30%) and low Lys (3%) content. The other bacterial samples have around 10% Arg and 35% Lys in this region.

4. DISCUSSION

As demonstrated above, there are indeed clear-cut differences between signal peptides from various species. In particular, eukaryotic signal peptides tend to have shorter n-, h- and c-regions than bacterial signal peptides, and among the latter, those from Gram-positive bacteria generally have longer n-, h- and c-regions than those from the Gram-negative *E. coli*. The net charge of the n-region is higher for the Gram-positive signal peptides.

Differences in the c-region may well be related to slightly different signal peptidase specificities in the different species. It is less obvious what may cause the differences in the n- and h-regions. A simple explanation would be that differences in the lipid composition of membranes play a role [9]. Thus, *E. coli* membranes contain more of the zwitterionic phosphatidylethanolamine and less of the negatively charged cardiolipin than, for example, *Bacillus* and *Streptomyces* [10,11].

Eukaryotic signal peptides have more hydrophobic h-regions than those of bacteria (more Leu and less Ala). The n-regions of the eukaryotic sequences contain less Lys (10–20% vs 30%) but approximately the same percentage of Arg (10%) as the prokaryotic ones.

Signal peptides from the Gram-positive *Strep-*

tomycetes are distinct from all other signal peptides in that their n-regions are much longer and contain much Arg (~30%) but almost no Lys (3%) residues ($p < 10^{-4}$). It is likely that these peculiarities result, at least in part, from the high G+C content of these bacteria (~70–75%) and the fact that lysine codons are GC-poor (AAA, AAG) whereas arginine codons are GC-rich (CGN, AGA, AGG). However, mature cytoplasmic and extracellular proteins from *Streptomyces* contain about the same percentages of Lys (3%) but only 6–7% Arg (not shown). Possibly, Arg can functionally replace Lys in signal peptides but not with equal ease in the mature parts of proteins.

Nevertheless, the excessive lengths of the Arg-rich n-regions from *Streptomyces* could be an indication that Arg may be in some sense less 'efficient' than Lys in signal peptide n-regions, making it necessary for these regions to be longer and/or more highly charged. We have tested this idea by comparing n-regions from eukaryotic signal peptides that contain only Arg but no Lys (108 sequences) with those containing only Lys but no Arg (100 sequences). Indeed, the former group tends to have longer (mean length 5.7 residues, $p < 10^{-4}$) but not more charged (mean net charge +1.2) n-regions than the latter (3.2 residues, +1.0) (fig.3). The Arg-containing n-regions also have a higher content of Pro than the Lys-containing ones (10% vs 1%, $p < 10^{-4}$); the same is true for n-regions from *Streptomyces* compared to n-regions from other bacterial species (12% vs 1%, $p < 10^{-3}$). Thus, when Arg is the only positively charged residue, it tends to be part of a rather long, possibly unstructured, n-region. Signal peptides from bacteria other than *Streptomyces* rarely have n-regions totally lacking Lys residues (none of these sequences listed in table 1 lack Lys in this region; also, out of the 59 *E. coli* sequences analyzed here, only 7 lack Lys in their n-region. Conversely, 8 out of 13 *Streptomyces* signal peptides contain Arg but no Lys; none contains Lys but no Arg).

A clear understanding of the differences in the c-regions is especially important when secretory proteins are expressed in foreign hosts. Although the (–3, –1)-rule seems to be valid for all organisms, more than one site compatible with this rule often exists in the vicinity of the normal cleavage site. It is therefore interesting to note that 'aberrant'

THIS PAGE BLANK (USPTO)

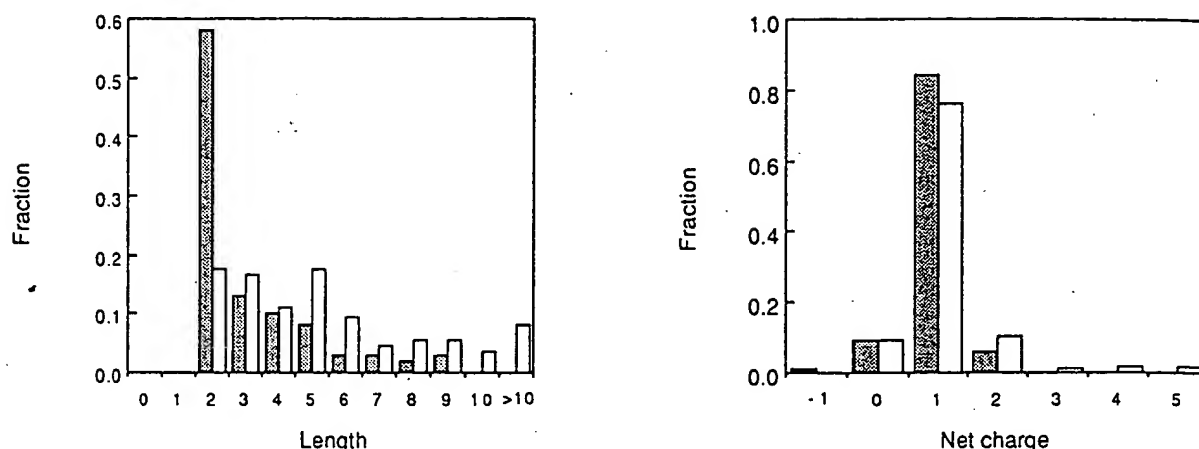


Fig.3. Length and net charge ($= n_{\text{Arg}} + n_{\text{Lys}} - n_{\text{Asp}} - n_{\text{Glu}}$) distributions for the n-regions of eukaryotic signal peptides that contain lysine (shaded bars) or arginine (unfilled bars) as the only positively charged residue.

cleavages have sometimes been found when signal peptides from Gram-positive bacteria have been expressed in *E. coli*. Thus, when α -amylase from *Bacillus stearothermophilus* was expressed and secreted by *E. coli*, a large fraction (40%) was incorrectly processed [12]. The aberrant cleavage took place three residues to the amino-terminal side of the normal cleavage site. This is what we would expect from the data presented here, namely that *E. coli* signal peptidase I seems to cleave preferentially six residues after the end of the h-region, whereas the corresponding proteases in Gram-positive bacteria seem to prefer a slightly longer distance from the h-region (7–9 residues).

Another example is provided by the major outer membrane lipoprotein from *E. coli*, which is normally cleaved by signal peptidase II. A mutant lipoprotein signal peptide that is not recognized by this peptidase is nevertheless cleaved by signal peptidase I when expressed in *E. coli* and yeast; however, the cleavage in yeast takes place four residues upstream of the *E. coli* cleavage site [13]. Generalizing from these examples, we anticipate that some eukaryotic signal peptides will be found to be cleaved downstream of their natural cleavage site when expressed in bacterial, and in particular Gram-positive, hosts.

A phenomenon of considerable biotechnological interest has recently been shown to be caused by the use of a heterologous signal peptide [19]. When the signal peptide from staphylococcal protein A

was used to secrete proteins in *E. coli*, the morphology of the host cell was affected and periplasmic proteins leaked out to the growth medium [15,19]. A similar 'leaky' phenotype of the host cell has also been observed when other Gram-positive proteins have been expressed and secreted in *E. coli* ([14,16–18]; Nygren, P.-Å., personal communication). Whether this effect was in all cases caused by the foreign signal peptide is unknown. In the light of the results presented above it is, however, tempting to suggest that this phenomenon is caused by a 'mismatch' between the signal peptide and the *E. coli* translocation machinery, resulting in a pleiotropic secretion defect with secondary effects on the structure of the outer membrane.

Acknowledgement: This work was supported by grants from the Swedish Natural Sciences Research Council.

REFERENCES

- [1] Von Heijne, G. (1985) *J. Mol. Biol.* 184, 99–105.
- [2] Von Heijne, G. (1986) *Nucleic Acids Res.* 14, 4683–4690.
- [3] Von Heijne, G. (1984) *EMBO J.* 3, 2315–2318.
- [4] Abrahmsén, L., Moks, T., Nilsson, B., Hellman, U. and Uhlen, M. (1985) *EMBO J.* 4, 3901–3906.
- [5] Smith, R.A., Duncan, M.J. and Moir, D.T. (1985) *Science* 229, 1219–1224.
- [6] Chang, C.N., Rey, M., Bochner, B., Heynecker, H. and Gray, G. (1987) *Gene* 55, 189–196.
- [7] Von Heijne, G. (1987) *Protein Seq. Data Anal.* 1, 41–42.

THIS PAGE BLANK (USPTO)

- [8] Levitt, M. (1978) *Biochemistry* 17, 4277-4285.
- [9] De Vrije, T., De Swart, R.L., Dowhan, W., Tommassen, J. and De Kruijff, B. (1988) *Nature* 334, 173-175.
- [10] Datta, D.B. (1987) *A Comprehensive Introduction to Membrane Biochemistry*, pp.72, Floral Publishing, Madison.
- [11] Sternberg, B., Gumpert, J., Reinhardt, G. and Gawrisch, K. (1987) *Biochim. Biophys. Acta* 898, 223-230.
- [12] Suominen, I., Karp, M., Lautamo, J., Knowles, J. and Mantsälä, P. (1987) *Extracellular Enzymes of Microorganisms* (Chaloupka, J. and Krumphanz, V. eds) pp.129-137, Plenum, New York.
- [13] Pines, O., Lunn, C.A. and Inouye, M. (1988) *Mol. Microbiol.* 2, 209-217.
- [14] Götz, F., Popp, F., Korn, E. and Schleifer, K.H. (1985) *Nucleic Acids Res.* 13, 5895-5906.
- [15] Abrahmsén, L., Moks, T., Nilsson, B. and Uhlén, M. (1986) *Nucleic Acids Res.* 14, 7487-7500.
- [16] Malke, H. and Ferretti, J.J. (1984) *Proc. Natl. Acad. Sci. USA* 81, 3557-3561.
- [17] Recsei, P.A., Gruss, A.D. and Novick, R.P. (1987) *Proc. Natl. Acad. Sci. USA* 84, 1127-1131.
- [18] Murphy, N., McConnell, D.J. and Cantwell, B.A. (1984) *Nucleic Acids Res.* 12, 5355-5367.
- [19] Abrahmsén, L., Nygren, P.-Å., Uhlén, M., Anderson, S. and Nilsson, B. (1988) submitted.
- [20] Yamazaki, H., Ohmura, K., Nakayama, A., Takeichi, Y., Ootai, K., Yamasaki, M., Tamura, G. and Yamane, K. (1983) *J. Bacteriol.* 156, 327-337.
- [21] Takkinen, K., Pettersson, R.F., Kalkkinen, N., Palva, I., Söderlund, H. and Kääriäinen, L. (1983) *J. Biol. Chem.* 258, 1007-1013.
- [22] Hussain, M., Carlino, A., Madonna, M.J. and Lampen, J.O. (1985) *J. Bacteriol.* 164, 223-229.
- [23] Tsuboi, A., Uchichi, R., Adachi, T., Sasaki, T., Hayakawa, S., Yamagata, Y., Tsukagoshi, N. and Udaka, S. (1988) *J. Bacteriol.* 170, 935-945.
- [24] Tsukamoto, A., Kimura, K., Ishii, Y., Takano, T. and Yamane, K. (1988) *Biochem. Biophys. Res. Commun.* 151, 25-31.
- [25] Sibakov, M. and Palva, I. (1984) *Eur. J. Biochem.* 145, 567-572.
- [26] Kimura, K., Kataoka, S., Ishii, Y., Takano, T. and Yamane, K. (1987) *J. Bacteriol.* 169, 4399-4402.
- [27] Takano, T., Fukuda, M., Monma, M., Kobayashi, S., Kainuma, K. and Yamane, K. (1986) *J. Bacteriol.* 166, 1118-1122.
- [28] Fukumori, F., Kudo, T., Narahashi, Y. and Horikoshi, K. (1986) *J. Gen. Microbiol.* 132, 2329-2335.
- [29] Nakajima, R., Imanaka, T. and Aiba, S. (1985) *J. Bacteriol.* 163, 401-406.
- [30] Steinmetz, M., Le Coq, D., Aymerich, S., Gonzy-Treboul, G. and Gay, P. (1985) *Mol. Gen. Genet.* 200, 220-228.
- [31] Tsuboi, A., Uchihi, R., Tabata, R., Takahashi, Y., Hashiba, H., Sasaki, T., Yamagata, H., Tsukagoshi, N. and Udaka, S. (1986) *J. Bacteriol.* 168, 365-373.
- [32] Fukusaki, E., Panbangred, W., Shinmyo, A. and Okada, H. (1984) *FEBS Lett.* 171, 197-201.
- [33] Wong, S.-L., Price, C.W., Goldfarb, D.S. and Doi, R.H. (1984) *Proc. Natl. Acad. Sci. USA* 81, 1184-1188.
- [34] Wong, S.-L. and Doi, R.H. (1986) *J. Biol. Chem.* 261, 10176-10181.
- [35] Kawazu, T., Nakanishi, Y., Uozumi, N., Sasaki, T., Yamagata, H., Tsukagoshi, N. and Udaka, S. (1987) *J. Bacteriol.* 169, 1564-1570.
- [36] Yamada, A., Tsukagoshi, N., Udaka, S., Sasaki, T., Nakamura, S., Little, C., Tomita, M. and Ikezawa, H. (1988) *Eur. J. Biochem.* 175, 213.
- [37] Hussain, M., Pastor, F.I.J. and Lampen, J.O. (1987) *J. Bacteriol.* 169, 579-586.
- [38] Shimada, H., Honjo, M., Mita, I., Nakayama, A., Akaoka, A., Manabe, K. and Furutani, Y. (1985) *J. Biotechnol.* 2, 75-85.
- [39] Metz, R.J., Allen, L.N., Cao, T.M. and Zeman, N.W. (1988) *Nucleic Acids Res.* 16, 5203.
- [40] Schörgendorfer, K., Schwab, H. and Lafferty, R.M. (1987) *Nucleic Acids Res.* 15, 9606.
- [41] Nielsen, J.B.K. and Lampen, J.O. (1982) *J. Biol. Chem.* 257, 4490-4495.
- [42] Wang, W., Mezes, P.S.F., Yang, Y.Q., Blacher, R.W. and Lampen, J.O. (1985) *J. Bacteriol.* 163, 487-492.
- [43] Lim, H.M., Pène, J.J. and Shaw, R.W. (1988) *J. Bacteriol.* 170, 2873-2878.
- [44] MacKay, R.M., Lo, A., Willick, G., Zuker, M., Baird, S., Dove, M., Moranelli, F. and Seligy, V. (1986) *Nucleic Acids Res.* 14, 9159-9170.
- [45] Paddon, C.J. and Hartley, R.W. (1986) *Gene* 40, 231-239.
- [46] Jacobs, M., Eliasson, M., Uhlén, M. and Flock, J.-I. (1985) *Nucleic Acids Res.* 13, 8913-8926.
- [47] Takagi, M., Imanaka, T. and Aiba, S. (1985) *J. Bacteriol.* 163, 824-831.
- [48] Murphy, N., McConnell, D.J. and Cantwell, B.A. (1984) *Nucleic Acids Res.* 12, 5355-5367.
- [49] Shortle, D. (1983) *Gene* 22, 181-189.
- [50] Jackson, M.P. and Iandolo, J.J. (1986) *J. Bacteriol.* 167, 726-728.
- [51] Löfdahl, S., Guss, B., Uhlén, M., Philipson, L. and Lindberg, M. (1983) *Proc. Natl. Acad. Sci. USA* 80, 697-701.
- [52] Berley, M.J. and Mekalanos, J.J. (1988) *J. Bacteriol.* 170, 34-41.
- [53] Gray, G.S. and Kehoe, M. (1984) *Infect. Immun.* 46, 615-618.
- [54] Bohach, G.A. and Schlievert, P.M. (1987) *Mol. Gen. Genet.* 209, 15-20.
- [55] Liebl, W. and Götz, F. (1986) *Mol. Gen. Genet.* 204, 166-173.
- [56] Carmona, C. and Gray, G.L. (1987) *Nucleic Acids Res.* 15, 6757.
- [57] Recsei, P.A., Gruss, A.D. and Novick, R. (1987) *Proc. Natl. Acad. Sci. USA* 84, 1127-1131.
- [58] Sako, T. and Tsuchida, N. (1983) *Nucleic Acids Res.* 11, 7679-7693.
- [59] Lee, C.Y. and Iandolo, J.J. (1986) *J. Bacteriol.* 166, 385-391.
- [60] Sakurai, S., Suzuki, H. and Kondo, I. (1988) *J. Gen. Microbiol.* 134, 711-714.
- [61] Johnson, L.P., L'Italien, J.J. and Schlievert, P.M. (1986) *Mol. Gen. Genet.* 203, 354-356.

THIS PAGE BLANK (USPTO)

- [62] Fahnestock, S.R., Alexander, P., Nagle, J. and Filpula, D. (1986) *J. Bacteriol.* 167, 870-880.
- [63] Olsson, A., Eliasson, M., Guss, B., Nilsson, B., Hellman, U., Lindberg, M. and Uhlén, M. (1987) *Eur. J. Biochem.* 168, 319-324.
- [64] Hollingshead, S.K., Fischetti, V.A. and Scott, J.R. (1986) *J. Biol. Chem.* 261, 1677-1686.
- [65] Malke et al. (1985) *Gene* 34, 357-362.
- [66] Shiroza, T. and Kuramitsu, H.K. (1988) *J. Bacteriol.* 170, 810-816.
- [67] Kehoe, M.A., Miller, L., Walker, J.A. and Boulnois, G.J. (1987) *Infect. Immun.* 55, 3228-3232.
- [68] Long, C.M., Virolle, M.-J., Chang, S.-Y., Chang, S. and Bibb, M.J. (1987) *J. Bacteriol.* 169, 5745-5754.
- [69] Guijarro, J., Santamaria, R., Schauer, A. and Losick, R. (1988) *J. Bacteriol.* 170, 1895-1901.
- [70] Argarona, C.E., Kuntz, I.D., Birken, S., Axel, R. and Cantor, C.R. (1986) *Nucleic Acids Res.* 14, 1871-1882.
- [71] Buttner, M.J., Fearnley, I.M. and Bibb, M.J. (1987) *Mol. Gen. Genet.* 209, 101-109.
- [72] Duez, C., Piron-Fraipont, C., Joris, B., Dusart, J., Urdea, M.S., Martial, J.A., Frère, J.-M. and Ghuyssen, J.-M. (1987) *Eur. J. Biochem.* 162, 509-518.
- [73] Hoshiko, S., Makabe, O., Nojiri, C., Katsumuta, K., Satoh, E. and Nagaoka, K. (1987) *J. Bacteriol.* 196, 1029-1036.
- [74] Nakai, R., Horinuchi, S. and Beppu, T. (1988) *Gene* 65, 229-238.
- [75] Robbins, P.W., Trimble, R.B., Wirth, D.F., Hering, C., Maley, F., Maley, G.F., Das, R., Gibson, B.W., Royal, N. and Biemann, K. (1984) *J. Biol. Chem.* 259, 7577-7583.
- [76] Dehottay, P., Dusart, J., De Meester, F., Joris, B., Van Beeumen, J., Epricum, T., Frère, J.-M. and Ghuyssen, J.-M. (1987) *Eur. J. Biochem.* 166, 345-350.
- [77] Eckhardt, T., Strickler, J., Gorniak, L., Burnett, W.V. and Fare, L.R. (1987) *J. Bacteriol.* 169, 4249-4256.
- [78] Henderson, G., Krygsman, P., Liu, C.J., Davey, C.C. and Malek, L.T. (1987) *J. Bacteriol.* 169, 3778-3784.
- [79] Burnett, W.V., Henner, J. and Eckhardt, T. (1987) *Nucleic Acids Res.* 15, 3926.
- [80] Liljeström, P.L. (1985) *Nucleic Acids Res.* 13, 7257-7268.
- [81] Arima, K., Oshima, T., Kubota, I., Nakamura, N., Mizunaga, T. and Toh-e, A. (1983) *Nucleic Acids Res.* 11, 1657-1672.
- [82] Bird, P., Gething, M.-J. and Sambrook, J. (1987) *J. Cell Biol.* 105, 2905-2914.
- [83] Stark, M.J.R. and Boyd, A. (1986) *EMBO J.* 5, 1995-2002.
- [84] Taussig, R. and Carlson, M. (1983) *Nucleic Acids Res.* 11, 1943-1954.
- [85] Kurjan, J. and Herskowitz, I. (1982) *Cell* 30, 933-943.
- [86] Waters, M.G., Evans, E.A. and Blobel, G. (1988) *J. Biol. Chem.* 263, 6209-6214.
- [87] Ammerer, G., Hunter, C.P., Rothman, J.H., Saari, G.C., Valls, L.A. and Stevens, T.H. (1986) *Mol. Cell Biol.* 6, 2490-2499.
- [88] Singh, A., Chen, E.Y., Lugovoy, J.M., Chang, C.N., Hitzeman, R.A. and Seeburg, P.H. (1983) *Nucleic Acids Res.* 11, 4049-4063.
- [89] MacKay, V.L., Welch, S.K., Insley, M.Y., Manney, T.R., Holly, J., Saari, G.C. and Parker, M.L. (1988) *Proc. Natl. Acad. Sci. USA* 85, 55-59.
- [90] Skipper, N., Thomas, D.Y. and Lau, P.C.K. (1984) *EMBO J.* 3, 107-111.
- [91] Yamashita, I., Suzuki, K. and Fukui, S. (1985) *J. Bacteriol.* 161, 567-573.
- [92] Stark, M.J.R., Mileham, A.J., Romanos, M.A. and Boyd, A. (1987) *Nucleic Acids Res.* 15, 6011-6030.

THIS PAGE BLANK (USPTO)

SIGPEP: a sequence database for secretory signal peptides

Gunnar von Heijne

Research Group for Theoretical Biophysics, Department of Theoretical Physics, Royal Institute of Technology,
S-100 44 Stockholm, Sweden

Many proteins, made by both prokaryotic and eukaryotic cells, are secreted across the cell membrane to the extracellular space. Such proteins are generally made with cleavable N-terminal extensions called variably signal sequences, signal peptides, or leader peptides. The workings of the protein export machinery and the role of the signal peptide have been extensively studied over the past 10-15 years, and considerable progress has been made both in defining the biochemical components involved in export and the structural determinants inherent in the signal peptide (see [1, 2] for reviews).

In this note, I describe a database collection, Sequence Database for Secretory Signal Peptides (SIGPEP), of both prokaryotic and eukaryotic signal peptides that has been used extensively in the study of their amino acid sequence characteristics [2-4]. The database is managed using the Appleworks program running on Apple IIe, Apple III, and Macintosh Plus microcomputers and is available either in Appleworks or Microsoft File format on standard floppy diskettes sent to the author.

A typical entry is shown in Fig. 1. The asterisk in the sequence indicates the point of cleavage between the signal

peptide and the mature protein, when this has been determined. At the time of writing, the database has 94 prokaryotic and 503 eukaryotic entries, roughly two-thirds of which have known cleavage sites.

The most prominent eukaryotic species represented in the database are: human (147 entries), mouse and rat (both 67 entries), and bovine (20 entries); among prokaryotic sequences *Escherichia coli* (43 entries) and *Bacillus* (19 sequences) are the most abundant. Thirty-five plant and 33 viral sequences are also included. Surprisingly, there are only eight yeast sequences in the database so far.

Many of these signal peptides can be extracted, with some work, from the main protein sequence data collections, such as the National Biomedical Research Foundation Protein Sequence Databank [5]. However, the large databases are plagued with a backlog of sequences not yet entered, going back some 1-2 years [Science (1986) 232:1599]. As is clear from Fig. 2, this means that as much as 50% of the data available in the literature at any one time cannot be retrieved except from specialized collections, such as SIGPEP. In fact, SIGPEP itself is probably not even reasonably complete for 1985, since the overall expo-

Species	guinea pig
Protein type	plasma hormone
Protein	insulin
Notes	cDNA, n.b. G in -3, not N as reported in PNAS 81,5046
Author	Watt, V.M.
Journal	J. Biol. Chem.
Year	1985
Volume	260
Pages	10926-10929
Sequence	MALWMHLLTLVLALLALWGPNTGQA*FVSRHLCGSN

Fig. 1. A typical Sequence Database for Secretory Signal Peptides (SIGPEP) entry. The asterisk indicates the position of the cleavage site between the signal peptide and the mature protein

THIS PAGE BLANK (USPTO)

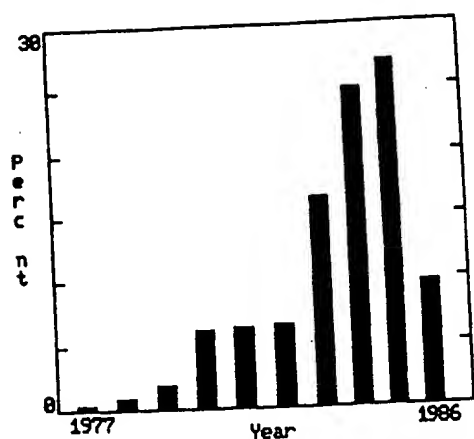


Fig. 2. Year-by-year distribution of SIGPEP entries as of July 1986

nential growth of sequence data would lead one to expect a larger number of entries from that year. Nevertheless, SIGPEP is probably the largest collection of signal peptides available.

References

1. Hortsch M, Meyer DI (1986) *Int Rev Cytol* 102:215-242
2. von Heijne G (1985) *Curr Top Membr Transp* 24:151-179
3. von Heijne G (1985) *J Mol Biol* 184:99-105
4. von Heijne G (1986) *Nucleic Acids Res* 14:4683-4690
5. Orcutt BC, George DG, Dayhoff MO (1983) *Annu Rev Biophys Bioeng* 12:419-441

Received July 1, 1986 / Accepted December 12, 1986

THIS PAGE BLANK (USPTO)

Net N-C Charge Imbalance May be Important for Signal Sequence Function in Bacteria

Gunnar von Heijne

Research Group for Theoretical Biophysics
Department of Theoretical Physics
Royal Institute of Technology
S-100 44 Stockholm, Sweden

(Received 3 March 1986, and in revised form 4 July 1986)

The net charge distribution in a region around the signal sequence cleavage site has been calculated for samples of 41 prokaryotic and 165 eukaryotic exported proteins. The results show that prokaryotic proteins in particular have a markedly higher incidence of acidic than of basic residues in this region. The possibility that a "dipolar" structure with a positive net charge difference between the N and C-terminal regions is important for signal sequence function in bacteria is suggested, and invoked to rationalize a number of known export-defective signal sequence mutations.

1. Introduction

Protein export in both prokaryotic and eukaryotic cells is initiated by an N-terminal signal sequence that routes the protein into the secretory pathway. The signal sequence has been studied intensively by biochemical, genetic and theoretical means over the past decade (Benson *et al.*, 1985; Oliver, 1985; von Heijne, 1985a), and at least three structurally and functionally distinct regions have been defined: a positively charged N-terminal region (n-region), typically one to five residues long, a seven to 15 residue long unbroken hydrophobic core region (h-region), and a typically five to six residue long more polar C-terminal region (c-region; von Heijne, 1985b). Generally, the introduction of charged or polar residues in the h-region often greatly reduces export, whereas alterations in the c-region primarily influence the efficiency of cleavage of the signal sequence from the mature protein (von Heijne, 1984a). Parts of the mature sequence can affect export, at least in bacteria (Bankaitis *et al.*, 1985). The precise structure-function relationships behind the observed effects are poorly understood.

In this paper, I wish to draw attention to a new sequence pattern that is particularly strong in bacterial signal sequences, and that might help explain in more detail the observed effects of charged residues in various parts of these peptides and in the adjoining region of the mature proteins.

2. Methods

A total of 41 prokaryotic and 165 eukaryotic signal sequences with known cleavage sites, chosen from my

collection of published sequences (currently counting some 80 prokaryotic and 430 eukaryotic entries) so as not to include any highly homologous ones, have been analysed with regard to net n-region charge, net c-region charge (including from 0 to 10 residues from the mature sequences), and net N-C charge difference (counting Arg and Lys as +1, Asp and Glu as -1). Also, the full prokaryotic collection including homologous sequences and sequences with cleavage sites predicted according to the "(-3, -1)-rule" (von Heijne, 1983, 1986; a total of 81 entries) has been analysed similarly.

As controls, theoretical net charge distributions expected for mature sequences with overall amino acid frequencies as in soluble proteins in general (45 exported bacterial proteins totalling 15,258 residues extracted from the National Biomedical Research Foundation Protein Sequence Database, Release 7.0, with $f_{Asp+Glu} = 0.118$ and $f_{Arg+Lys} = 0.100$; and eukaryotic amino acid frequencies as given by Klapper (1977), i.e. $f_{Asp+Glu} = 0.117$ and $f_{Arg+Lys} = 0.117$) have been derived; e.g. the expected frequencies of all bacterial 5-residue segments with a net charge of -5 and -4 are, respectively, $(0.118)^5$ and $5 \times 0.782 \times (0.118)^4$, etc. The expected distributions of net N-C charge differences have been calculated by weighting the observed net n-region charge distributions with the theoretical C-terminal charge distributions calculated above. Mean values have been compared using a 2-sided *t*-test.

3. Results

Both charged and hydrophobic residues have been implicated in the proper functioning of the signal sequence. The importance of an intact h-region has been demonstrated repeatedly (Oliver, 1985). In prokaryotes, a net n-region charge ≥ 0 seems to be required (Vlasuk *et al.*, 1983) and a

THIS PAGE BLANK (USPTO)

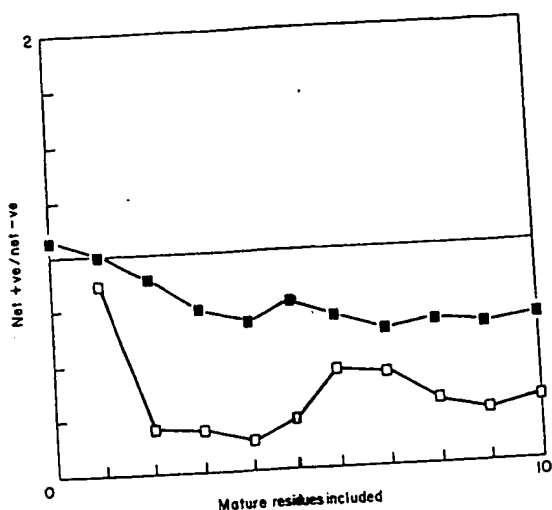


Figure 1. Quotient between the number of sequences with net positive and net negative C-terminal charge including all residues in the c-region plus an additional zero to 10 residues from the mature protein. Filled squares: eukaryotic proteins ($N = 165$); open squares: prokaryotic proteins ($N = 41$).

similar, although perhaps slightly less strict, n-region charge distribution has been observed in eukaryotes (von Heijne, 1984b). Less attention has been paid so far to the possible role of charged residues in the c-region and the early parts of the mature protein.

The number of known signal sequences has now increased sufficiently to make rather detailed statistical studies feasible even for prokaryotic proteins. In particular, patterns of charged residues in the region around the cleavage site less obvious to the eye than the nearly universal occurrence of basic residues in the n-region can now be explored and, as shown below, the distribution of charged residues around the C-terminal end of prokaryotic signal sequences is definitely non-random.

Indeed, the incidence of negatively and positively charged residues in the c-region and the first few positions in the mature sequences is in fact almost the inverse of that in the signal sequence n-region: acidic residues abound and basic residues are scarce (Fig. 1), resulting in most cases in a negative net C-terminal charge, especially evident in the prokaryotic sample (Fig. 2). (The mean net C-terminal charge calculated for the distribution in Fig. 2 (-0.66) is significantly different from the value expected for the random distribution (-0.09 ; $P < 0.001$), whereas the corresponding eukaryotic distribution has only a marginally significant lower mean net charge (-0.22) than the random one (± 0.0 ; $P < 0.05$.) In particular, position +2 in the mature prokaryotic sequences is strongly biased towards negatively charged residues, with eight Asp (20%), nine Glu (22%), and not a single Arg or Lys observed in the sample. Altogether, only five sequences, *versus* an expected 11 for a sample with

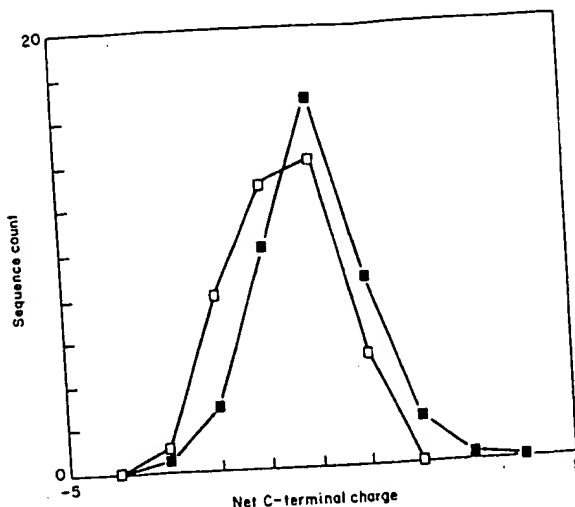


Figure 2. Theoretical (filled squares, see Methods), and observed (open squares) net C-terminal charge distribution for the prokaryotic signal peptides ($N = 41$), including all residues in the c-region plus 5 residues from the mature protein.

amino acid frequencies as in soluble proteins in general (see Methods), have a positive net C-terminal charge when five residues from the mature chains are included in the calculation (Fig. 2).

This skewed distribution of charged residues in the C-terminal region adds to the well-known bias towards Arg and Lys residues in the N-terminal region and gives the typical prokaryotic signal sequence a highly "dipolar" appearance. In fact, even if we assume that the initiator f-Met is retained and hence does not contribute any charge (von Heijne, 1984b), 40 out of the 41 sequences in the prokaryotic sample have a net N-C charge difference $\Delta_{N-C} > 0$ (calculated with 5 mature residues included; this choice is not very critical, cf. Fig. 1, and is used throughout), and the remaining one, *phoA*, has $\Delta_{N-C} = 0$. Given the theoretical net C-terminal charge distribution in Figure 2 and the observed net N-terminal charge distribution for the prokaryotic sample, one would expect to find four sequences with $\Delta_{N-C} \leq 0$ in a sample with normal amino acid frequencies in the C-terminal region. For the full 81-sequence prokaryotic sample (see Methods), a similar calculation gives eight expected sequences; again, *phoA* with $\Delta_{N-C} = 0$ is the only one found.

4. Discussion

The observed bias towards negatively charged Asp and Glu residues in the immediate N-terminal region of mature exported prokaryotic proteins, being particularly marked in position +2 relative to the signal sequence cleavage site, may be related primarily to the cleavage reaction. However, it has been shown that position +1 is highly enriched for Ala in prokaryotic signal peptides (von Heijne,

THIS PAGE BLANK (USPTO)

phoA + nuclease		
wt	MKQSTIALALLPLLFTPVTKA*RTPEMPVLEN...	100%
pFOG403	-----*ATSTMLKEZ...	100%
pFOG406	-----RATSTMLHK...	0%
β -Lactamase + insulin		
wt	MSIQHFRVALIPFFAAFLPVFA*HPETLVKVD...	100%
125/-7	-----*HPLQEPKPA...	50%
124/-7	-----*HRCSEPKPAQ...	50%
pJW2172	-----ALLALWGDPAAAA*FVKQHLGG...	100%
β -Lactamase + tpi		
PTG2	MRIQHFRVALIPFFAAFLPVFG*HPETLVKVD...	100%
PTSS1	-----MAPRKFVGG...	0%
PTSS2	-----PPRKFVCGH...	0%

Figure 3. Fusion proteins discussed in the text (Talmadge *et al.*, 1980; Chan *et al.*, 1981; Kadonaga *et al.*, 1984). An asterisk indicates the cleavage site; dashes show homologous residues. wt, wild-type. The efficiency of export is given at the right.

1984a), possibly reflecting some functional constraint on a free N terminus (Flinta *et al.*, 1986); thus, position +2 is the obvious choice for introducing an Asp or Glu as close to the signal sequence C terminus as possible, even for reasons that have nothing to do with cleavage (in prokaryotes but not in eukaryotes, charged residues seem to be avoided in the c-region proper).

A more intriguing possibility than an effect on cleavage is that the prokaryotic export machinery may require a net positive charge difference between the N and C-terminal regions of the signal sequence. If so, we should expect to find non-functional signal sequence mutations that destroy such an imbalance and, indeed, a number of good candidates can be found in the literature (Fig. 3).

As noted above, phoA from *Escherichia coli* is the only example found so far with $\Delta_{N-C} = 0$: it has a single Lys in the n-region, and another one in position -2, followed by an Arg in +1 and a Glu in +4. This signal sequence was recently fused to the mature part of *Staphylococcus aureus* nuclease, and was shown to induce export of a correctly cleaved product in *E. coli* (Liss *et al.*, 1985). When a single Arg (as found in the normal phoA) was introduced between the signal sequence and the nuclease, however, export was completely inhibited. A possible explanation based on the results presented here might be the absence in the beginning of the nuclease sequence of any negatively charged residues to counteract the negative Δ_{N-C} value of the latter construction.

A similar case, where the β -lactamase signal sequence was fused to various parts of preproinsulin

and expressed in *E. coli*, has been described (Talmadge *et al.*, 1980; Chan *et al.*, 1981; see Fig. 3). Here, two fusions with intact signal sequences but with $\Delta_{N-C} = 0$ are secreted poorly (50%), whereas all those with $\Delta_{N-C} > 0$ function well, even when half of the h-region is provided by the insulin signal sequence (pJW2172).

It is even possible that the failure to secrete chicken triose phosphate isomerase (*tpi*, a non-exported enzyme) from *E. coli* with the aid of a modified β -lactamase signal sequence (Kadonaga *et al.*, 1984) may be due, at least in part, to the occurrence of one Arg and one Lys among the first five *tpi* residues ($\Delta_{N-C} = 0$).

Rather than increasing the net C-terminal charge, decreasing the net N-terminal charge should similarly have an adverse effect on export. Results on the secretion of *E. coli* lipoprotein are consistent with this idea: when the N-terminal net charge is reduced such that $\Delta_{N-C} < 0$, export (as well as synthesis) is severely inhibited (Vlasuk *et al.*, 1983). A *lamB* mutant with an Arg \rightarrow Ser replacement in the n-region, however, makes less of the protein, but does not appear to have a defect in secretion *per se* (Hall *et al.*, 1983); in this case, $\Delta_{N-C} = +2$.

Finally, a number of mutations and reversions in and around the signal sequence h-region have been isolated over the years (Fig. 4). From the present perspective, it is noteworthy that positively charged residues seem to reduce export more efficiently when placed towards the C-terminal end of the h-region (e.g. *malE* Met_{-8,-9} \rightarrow Arg *versus* Leu₋₁₇ \rightarrow Arg), whereas the opposite seems to hold for negatively charged residues (*malE* Ala₋₁₆ \rightarrow Glu *versus* lpp Gly₋₇ \rightarrow Asp). The notable observation that both Asp and Arg are tolerated in position -9 in *lamB* when Glu₋₁₀ and Arg₋₇ are not might be explained if we assume that the nine-residue long non-polar stretch -18 to -10 can function as an h-region (with the correct charge polarity) in the Asp₋₉ mutant, as can the non-polar stretch between -8 and +1 (again with correct charge polarity) in the Arg₋₉ mutant. A similar explanation may be advanced for the *malE* 18-1 and 19-1 revertants shown: both the 18-1 Ser \rightarrow Phe and the 19-1 Lys \rightarrow (Thr, Asn) reversions may create marginally functional h-regions with correct charge polarity in the former c-region of the signal sequence. It is interesting to note, finally, that of all the *malE* revertants tested to date (except for reversions back to wild-type) only one, where a possible h-region is created on the N-terminal side of the Ala₋₁₃ \rightarrow Glu mutation through an Arg₋₁₉ \rightarrow Leu replacement and thus giving rise to correct charge polarity, regains wild-type kinetics in export and processing (Ryan *et al.*, 1986).

It should be stressed, however, that many of the point mutations and deletions that have been isolated do not directly involve charged residues; in these cases, perturbations of the secondary structure of the signal peptide seems the most likely explanation (Briggs *et al.*, 1985).

In addition to the observed preference for

THIS PAGE BLANK (USPTO)

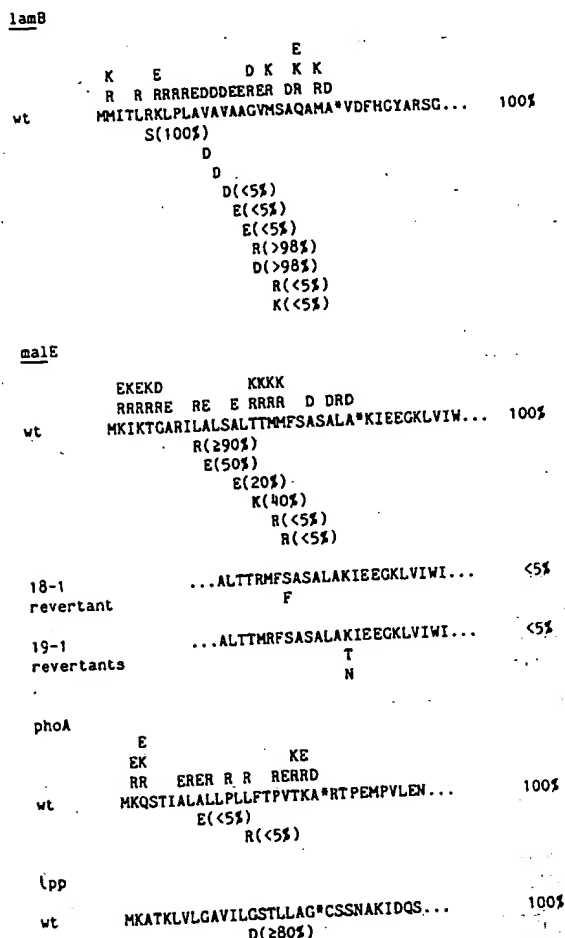


Figure 4. A summary of characterized signal sequence mutations involving substitutions of uncharged with charged residues (from a compilation by Oliver, 1985). The efficiency of export is indicated at the right. Above the wild-type sequences (wt), all theoretically possible mutations to charged residues involving only single-nucleotide changes are shown. An asterisk indicates the cleavage site.

positively charged residues towards the N terminus and negatively charged ones towards the C terminus of prokaryotic signal peptides, a considerable contribution to the net dipole moment of the peptide may be provided by the alignment of the individual peptide-bond dipoles when in a helical conformation, amounting to approximately half an additional positive charge placed at the N terminus and half a negative charge at the C terminus (Hol, 1985).

Obviously, if in addition to a potentially helical h-region of sufficient length and hydrophobicity

(von Heijne, 1985b), a charge polarity of the proposed kind is an important element in the function of the signal sequence, various interactions between this dipolar structure and the membrane potential (negative inside, positive outside) can be envisioned. In this respect, it is perhaps significant that export of β -lactamase, with $\Delta_{N-C} = +2$, is more sensitive to reductions in the membrane potential induced by CCCP (a proton ionophore) than is export of the *livK*-protein, which has $\Delta_{N-C} = +4$ (Daniels *et al.*, 1981). The likely absence of a sizeable membrane potential across the endoplasmic reticulum (Weinstein *et al.*, 1982) may similarly explain the much weaker preference for negatively charged residues in the C-terminal region found above for the eukaryotic signal peptides.

References

- Bankaitis, V. A., Ryan, J. P., Rasmussen, B. A. & Bassford, P. J. (1985). *Curr. Topics Membr. Trans.* **24**, 105-150.
- Benson, S. A., Hall, M. N. & Silhavy, T. J. (1985). *Annu. Rev. Biochem.* **54**, 101-134.
- Briggs, M. S., Gierasch, L. M., Zlotnick, A., Lear, J. D. & DeGrado, W. F. (1985). *Science (Wash.)*, **228**, 1096-1099.
- Chan, S. J., Weiss, J., Konrad, M., White, T., Bahl, C., Yu, S.-D., Marks, D. & Steiner, D. F. (1981). *Proc. Nat. Acad. Sci., U.S.A.* **78**, 5401-5405.
- Daniels, C. J., Bole, D. G., Quay, S. C. & Oxender, D. L. (1981). *Proc. Nat. Acad. Sci., U.S.A.* **78**, 5396-5400.
- Flinta, C., Persson, B., Jörnvall, H. & von Heijne, G. (1986). *Eur. J. Biochem.* **154**, 193-196.
- Hall, M. N., Gabay, J. & Schwartz, M. (1983). *EMBO J.* **2**, 15-19.
- Hol, W. G. (1985). *Prog. Biophys. Mol. Biol.* **45**, 149-195.
- Kadonaga, J. T., Gutier, A. E., Straus, D. R., Charles, A. D., Edge, M. D. & Knowles, J. R. (1984). *J. Biol. Chem.* **259**, 2149-2154.
- Klapper, M. H. (1977). *Biochem. Biophys. Res. Commun.* **78**, 1018-1024.
- Liss, L. R., Johnson, B. L. & Oliver, D. B. (1985). *J. Bacteriol.* **164**, 925-928.
- Oliver, D. (1985). *Annu. Rev. Microbiol.* **39**, 615-648.
- Ryan, J. P., Duncan, M. C., Bankaitis, V. A. & Bassford, P. J. (1986). *J. Biol. Chem.* **261**, 3389-3395.
- Talmadge, K., Stahl, S. & Gilbert, W. (1980). *Proc. Nat. Acad. Sci., U.S.A.* **77**, 3369-3373.
- Vlasuk, G. P., Inouye, S., Ito, H., Itakura, K. & Inouye, M. (1983). *J. Biol. Chem.* **258**, 7141-7148.
- von Heijne, G. (1983). *Eur. J. Biochem.* **133**, 17-21.
- von Heijne, G. (1984a). *J. Mol. Biol.* **173**, 243-251.
- von Heijne, G. (1984b). *EMBO J.* **3**, 2315-2318.
- von Heijne, G. (1985a). *Curr. Topics Membr. Trans.* **24**, 151-179.
- von Heijne, G. (1985b). *J. Mol. Biol.* **184**, 99-105.
- von Heijne, G. (1986). *Nucl. Acids Res.* **14**, 4683-4690.
- Weinstein, J. N., Blumenthal, R., van Renswoude, J., Kempf, C. & Klausner, R. D. (1982). *J. Membr. Biol.* **66**, 203-212.

THIS PAGE BLANK (USPTO)

on the other hand. A possible serological relationship between the coat proteins of the three strains is being investigated. The observation that TCM-cDNA probes cross-hybridize with PEBV-RNA 2 (Figure 4) correlates well with an earlier report that a Dutch isolate of PEBV is serologically related to a TRV serotype I-II isolate (19). The results shown in Figure 4 indicate that, in their association with TCM-RNA 1 and PEBV-RNA 1, the coding sequences of TCM-RNA 2 are accommodated with different 5'- and 3'-terminal sequences. It is possible that the mechanism of copy-choice, as discussed in the previous paragraph, permits the attachment of different termini to the TCM-RNA 2 coding sequence, thus enabling the replication of this sequence by the replicases encoded by either TCM-RNA 1 or PEBV-RNA 1. For a more detailed insight in this matter, sequence studies on the PEBV genome are required.

ACKNOWLEDGMENTS

Thanks are due to Mr. P. van Zuidam and Mr. F. Th. Brederode for assistance in the sequence studies. This work was supported in part by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

*To whom correspondence should be addressed

REFERENCES

1. Harrison, B.D. and Robinson, D.J. (1978) *Adv. Virus Res.* 23, 25-77.
2. Robinson, D.J. and Harrison, B.D. (1985) *J. gen. Virol.* 66, 171-176.
3. Cornelissen, B.J.C., Linthorst, H.J.M., Brederode, F.Th. and Bol, J.F. (1986) *Nucleic Acids Res.* 14, 2157-2169.
4. Mayo, M.A., Fritsch, C. and Hirth, L. (1976) *Virology* 69, 408-415.
5. Fritsch, C., Mayo, M.A. and Hirth, L. (1977) *Virology* 77, 722-732.
6. Pelham, H.R.B. (1979) *Virology* 97, 256-265.
7. Boccardo, M., Hamilton, W.D.O. and Baulcombe, D.C. (1986) *EMBO J.* 5, 223-229.
8. Bergh, S.T., Koziel, M.G., Huang, S.-C., Thomas, R.A., Gilley, D.P. and Siegel, A. (1985) *Nucleic Acids Res.* 13, 8507-8518.
9. Bol, J.F., Brederode, F.Th., Janze, G.C. and Rauh, D.C. (1976) *Virology* 65, 1-15.
10. Messing, J. (1983) *Methods Enzymol.* 101, 20-78.
11. Kiény, M.P., Lathe, R. and Lecocq, J.P. (1983) *Gene* 26, 91-99.
12. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
13. Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) *Proc. Natl. Acad. Sci. USA* 80, 3963-3965.
14. Hennikoff, S. (1984) *Gene* 28, 351-359.
15. Koper-Zwarthoff, E.C., Lockard, R.E., Alzner-De Weerd, B., RajBhandary, U.L. and Bol, J.F. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5504-5508.
16. Peattie, D. and Gilbert, W. (1980) *Proc. Natl. Acad. Sci. USA* 77, 4679-4682.
17. Peattie, D. (1979) *Proc. Natl. Acad. Sci. USA* 76, 1760-1764.
18. Donis-Keller, H. (1977). *Nucleic Acids Res.* 4, 2527-2538.
19. Maat, D.Z. (1963) *Neth. J. Plant Pathol.* 69, 287-239.

A new method for predicting signal sequence cleavage sites

Gunnar von Heijne

Research Group for Theoretical Biophysics, Department of Theoretical Physics, Royal Institute of Technology, S-100 44 Stockholm, Sweden

Received 5 March 1986; Revised and Accepted 5 May 1986

ABSTRACT

A new method for identifying secretory signal sequences and for predicting the site of cleavage between a signal sequence and the mature exported protein is described. The predictive accuracy is estimated to be around 75-80% for both prokaryotic and eukaryotic proteins.

INTRODUCTION

The transient N-terminal signal sequence found on most secretory proteins serves to initiate export across the inner membrane (in prokaryotes) or the endoplasmic reticulum (in eukaryotes). Three structurally and, possibly, functionally distinct regions have been identified as the basic building-blocks of a secretory signal sequence: a basic N-terminal region (n-region), a central hydrophobic region (h-region), and a more polar C-terminal region (c-region) (1). The structural determinants for cleavage of the signal sequence from the mature protein once export is under way seems to reside in the n- and h-regions, with positions -3 and -1 relative to the cleavage site being the most important ones (2,3). Indeed, this "(-3,-1)-rule" has been used quite successfully to predict the most like site of cleavage directly from the primary sequence (2).

In view of the great interest in secretory proteins and the fact that most such proteins are known only from their DNA sequence, it is important to assess and, if possible, to improve upon the predictive accuracy of the original method. In this paper, I present a new scheme based on a weight-matrix approach that can be expected to give correct predictions about 75-80% of the time when applied to new sequences (both prokaryotic and eukaryotic). This represents a substantial gain over the old method, which is shown to be around 65% and 45% accurate for eukaryotic and prokaryotic proteins, respectively.

THIS PAGE BLANK (USPTO)

METHODS

161 eukaryotic and 36 prokaryotic non-homologous signal sequences with known cleavage sites were chosen from my collection of signal sequences totalling at the present time some 450 eukaryotic and 80 prokaryotic entries. The prokaryotic sample did not include any sequences known to be cleaved by the lipoprotein signal peptidase (signal peptidase II) (4).

Weight-matrices $W(a,i)$ (see below) were calculated from the observed amino acid counts in each position, $N(a,i)$, (i.e. the number of residues of type a in position i) with all sequences aligned from their known site of cleavage between positions -1 and $+1$, by first dividing all counts by their respective expected abundance in proteins in general, $\langle N(a) \rangle$ (Tables 1 & 2, last column), and then taking the natural logarithms of these quotients: $W(a,i) = \ln(N(a,i)/\langle N(a) \rangle)$. To correct for the limited size of the data base, all zero-elements in the amino acid count matrices were put equal to one before the division. Zero-counts in positions -3 and $+1$ were treated differently: they were also put equal to one, but then divided by the total number of sequences in the sample, N , rather than the expected number of residues, e.g. $W(a,-1) = \ln(1/N)$ if $N(a,-1) = 0$.

The most probable cleavage site was identified by scanning the sequence in question with the appropriate weight-matrix and summing the weights for each position, i.e. $S(i) = W(a_1-p,i-p) + W(a_1-p,i-p+1) + \dots + W(a_1-q,i-q)$ where the summation window extends from position $i-p$ to $i-q$. The predicted cleavage site j is the one with the highest S -value, $S(j) = \max[S(i); i=1-p, \dots, L-q]$, where L is the length of the sequence analyzed. As shown below, maximum predictive accuracy was obtained for $p=12$ and $q=2$.

RESULTS

The $(-3,-1)$ -rule

Based on previous statistics (2), acceptable cleavage sites were suggested to conform to the following rules: the residue in position -1 must be small, i.e. either Ala, Ser, Gly, Cys, Thr, or Gln; the residue in position -3 must not be aromatic (Phe, His, Tyr, Trp), charged (Asp, Glu, Lys, Arg), or large and polar (Asn, Gln). Further, it was suggested that Pro must be absent from positions -3 through $+1$. The new amino acid counts presented in Tables 1 & 2 are based on more than twice as many sequences; nevertheless, the $(-3,-1)$ -rule is seen to hold remarkably well. The only exceptions found to date among eukaryotic proteins are one sequence with Leu in -1 , one with Pro in -2 , and three with Pro in -1 . Thus, barring sequencing errors, we must

Table 1 Amino acid counts for eukaryotic signal sequences
The average composition (last column) is from Ref. (10)

	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1	+2	Expected
A	16	13	14	15	20	18	18	17	25	15	47	6	80	18	6	14.5
C	3	6	9	7	9	14	6	8	5	6	19	3	9	8	3	4.5
D	0	0	0	0	0	0	0	0	5	3	0	5	0	10	11	8.9
E	0	0	0	1	0	0	0	0	3	7	0	7	0	13	14	10.0
F	13	9	11	11	6	7	18	13	4	5	0	13	0	6	4	5.6
G	4	4	3	6	3	13	3	2	19	34	5	7	39	10	7	12.1
H	0	0	0	0	0	1	1	0	5	0	0	6	0	4	2	3.4
I	15	15	8	6	11	5	4	8	5	1	10	5	0	8	7	7.4
K	0	0	0	1	0	0	1	0	0	4	0	2	0	11	9	11.3
L	71	68	72	79	78	45	64	49	10	23	8	20	1	8	4	12.1
M	0	3	7	4	1	6	2	2	0	0	0	1	0	1	2	2.7
N	0	1	0	1	1	0	0	0	3	3	0	10	0	4	7	7.1
P	2	0	2	0	0	4	1	8	20	14	0	1	3	0	22	7.4
Q	0	0	0	1	0	6	1	0	10	8	0	18	3	19	10	6.3
R	2	0	0	0	0	1	0	0	7	4	0	15	0	12	9	7.6
S	9	3	8	6	13	10	15	16	26	11	23	17	20	15	10	11.4
T	2	10	5	4	5	13	7	7	12	6	17	8	6	3	10	9.7
V	20	25	15	18	13	15	11	27	0	12	32	3	0	8	17	11.1
W	4	3	3	1	1	2	6	3	1	3	0	9	0	2	0	1.8
Y	0	1	4	0	0	1	3	1	1	2	0	5	0	1	7	5.6

admit the possibility that residues other than the classical $(-3,-1)$ -kinds can be used in position -1 , but only when no better cleavage site is available in the vicinity (this is true for all five exceptions).

A few other points can also be made. First, the constraints on the prokaryotic sequences in the $(-3,-1)$ -region seem even stronger than for the eukaryotic ones: only Ala, Gly, Ser and Thr have been found in -1 , and only Ala, Gly, Leu, Ser, Thr, and Val in -3 . Second, Leu is abundant in the prokaryotic sample up to and including position -8 , but its incidence drops precipitously in position -7 , where it is replaced by the likewise hydrophobic but less strongly helix-inducing residues Val and Phe. Only from position -6 do we find predominantly polar residues. Finally, there is a notable imbalance between the basic residues Arg and Lys in the c -region of the eukaryotic signal sequences, with 26 Arg and only 6 Lys (Arg/Lys = 4.3). This is in sharp contrast to the n -region where Arg/Lys = 66/72 = 0.9 and to proteins in general where the expected ratio is 0.6 (Table 1, last column).

THIS PAGE BLANK (USPTO)

Table 2 Amino acid counts for prokaryotic signal sequences
The average composition (last column) is from Ref. (10)

	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1	+2	Expected
A	10	8	8	9	6	7	5	6	7	7	24	2	31	18	4	3.2
C	1	0	0	1	1	0	0	1	1	0	0	0	0	0	0	1.0
D	0	0	0	0	0	0	0	0	0	0	0	0	0	2	8	2.0
E	0	0	0	0	0	0	0	0	0	0	0	1	0	4	8	2.2
F	2	4	3	4	1	1	8	0	4	1	0	7	0	1	0	1.3
G	4	2	2	2	3	5	2	4	2	2	0	2	2	1	0	2.7
H	0	0	1	0	0	0	0	1	1	0	0	7	0	1	0	0.8
I	3	1	5	1	5	0	1	3	0	0	0	0	0	0	2	1.7
K	0	0	0	0	0	0	0	0	0	1	0	2	0	3	0	2.5
L	8	11	9	8	9	13	1	0	2	2	1	2	0	0	1	2.7
M	0	2	1	1	3	2	3	0	1	2	0	4	0	0	1	0.6
N	0	0	0	0	0	0	0	1	1	1	0	3	0	1	4	1.6
P	0	1	1	1	1	1	2	3	5	2	0	0	0	0	5	1.7
Q	0	0	0	0	0	0	0	0	2	2	0	3	0	0	1	1.4
R	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1.7
S	1	0	1	4	4	1	5	15	5	8	5	2	2	0	0	2.6
T	2	0	4	2	2	2	2	2	5	1	3	0	1	1	2	2.2
V	5	7	1	3	1	4	7	0	0	4	3	0	0	2	0	2.5
W	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.4
Y	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	1.3

Construction of weight-matrices

Weight-matrix methods have been used for a number of years to locate signals in nucleic acid sequences (see (5) for a thorough discussion). Their use for pattern recognition in protein sequences requires a larger data base (20 amino acids rather than 4 bases must be scored for in each position), but is no different in principle. Basically, one converts the observed number of each kind of residue in each position in a sample of aligned "signals" into a measure of the probability of finding that particular kind of residue in that particular position - the probability weight-matrix - by a suitable normalization. Then, any new sequence can be scanned by a moving window (looking up the respective probabilities in the weight-matrix and multiplying together for each position of the window) to get a measure of the fit to the sample used in the construction of the weight-matrix. The highest-scoring window-position is then taken as the prediction for the location of the signal, if the score is above some minimum value.

To score for possible signal sequence function, and to locate the most probable cleavage site in a putative signal sequence, weight-matrices for prokaryotic and eukaryotic signal sequences were constructed as follows. The raw amino acid counts for the two samples (Tables 1 & 2) were divided by the expected number $\langle N(a) \rangle$ of each kind of residue given amino acid frequencies as in soluble proteins in general (last columns). Except for positions -3 and -1 relative to the cleavage site, all matrix elements with zero counts were normalized as $1/\langle N(a) \rangle$. For positions -3 and -1, where there is good reason both from previous statistical and experimental studies to believe that only a subset of all residues are allowed (2,6), the more stringent normalization $1/N$ was used for the zero-count elements (where N is the total number of sequences in the sample). The final weight-matrix was obtained by taking natural logarithms of the normalized values, thus reducing the ensuing probability calculations to summations rather than multiplications of the weight-matrix elements.

Assessment of the predictive accuracy

When the two weight-matrices were used to predict the cleavage sites in the samples used in their construction, virtually all sites were correctly identified (87% in the eukaryotic sample, 100% in the prokaryotic sample). However, these sequences are at an advantage relative to sequences not included in the matrix: when correctly aligned with the weight-matrix, all residues in a sequence included in the weight-matrix sample will correspond to a count, and a spuriously high predictive accuracy will be found.

To avoid this problem, the eukaryotic sample was divided into 7 subsamples, each of 23 sequences. For each subsample, the remaining 138 sequences were used to construct a new weight-matrix, and this matrix was then applied to the subsample. Similarly, the prokaryotic sample was divided into 4 subsamples, each of 9 sequences. All subsequent calculations were carried out by summing the results for the subsamples.

Weight-matrices including positions -15 to +5 were first used to determine the effect of ignoring residues at either end in the predictions. It was found that positions -13 to +2 were sufficient to obtain maximal predictive accuracy (for the prokaryotic sample, positions -5 to +2 were sufficient but the full -13 to +2 range was used nevertheless); with this choice, 125 out of 161 eukaryotic and 32 out of 36 prokaryotic cleavage sites (78% and 89%) were correctly identified with a standard deviation of about $\pm 10\%$ in each case. For an additional 19 eukaryotic and 2 prokaryotic sequences, the correct site had the second-highest score. These values should

THIS PAGE BLANK (USPTO)

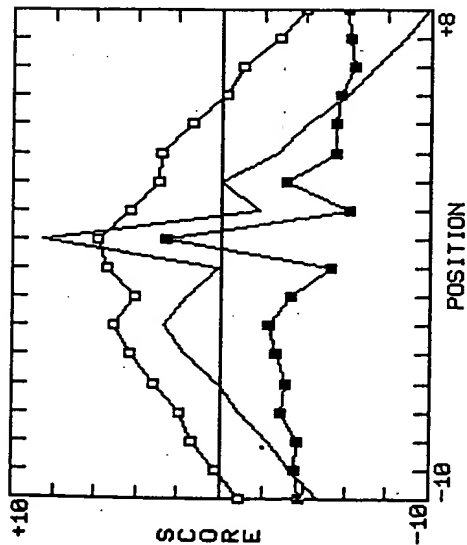


Figure 1 Average h- and c-region scores as a function of the position of the moving window. Open squares: h-region; solid squares: c-region; full line: total score.

be compared with the predictive accuracy of the older method (as implemented in a program kindly communicated by Dr. H.S. Ip, Rockefeller University). When this method was applied to the 121 sequences in the eukaryotic sample that were not included in the original statistics (2), 77/121 (64%) of the known cleavage sites were correctly identified, and only 17/36 (47%) of the prokaryotic ones were found.

With -13 to +2 weight-matrices, the contribution to the overall success from individual positions was also investigated. Only positions -3 and -1 had any strong impact; when one or the other was left out in the calculations the percentage of correctly identified eukaryotic sites dropped to 61% and 53%, respectively (81% and 69% for the prokaryotic sample).

As has been shown previously (1,7), residues -13 to -6 correspond to the h-region in the "average" eukaryotic signal sequence, residues -5 to -1 correspond to the c-region, and residues +1 and +2 seem to be selected such that few alternative cleavage sites should exist in the vicinity of the correct one (i.e. residues -5 to +2 can be included in an extended c-region). Thus, it is possible to calculate the scores for the h- and c-regions separately by summing the contributions from positions -13 to -6 and -5 to +2, respectively. As shown in Fig.1, the average h-region score for the eukaryotic sample increases slowly as the window is moved up to position -1 (the known cleavage site), and then decreases. The average c-region score

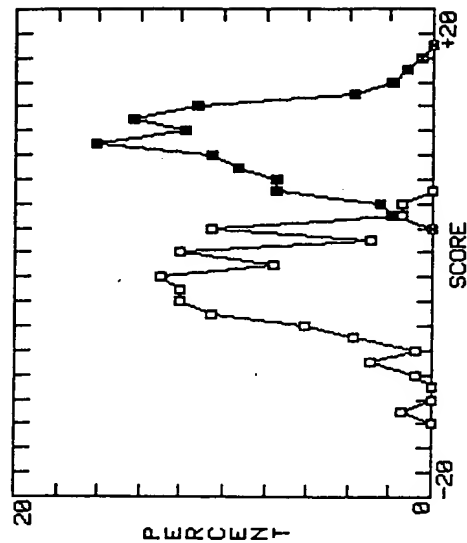


Figure 2 Distribution of maximum scores for signal sequences and cytosolic proteins. Open squares: cytosolic proteins; solid squares: signal sequences.

shows a more dramatic behaviour, with a pronounced peak in position -1 and troughs in positions -2 and +1, reflecting the match to the (-3,-1)-pattern and the tendency to have residues in position -2 that do not fit this pattern (see Tables 1 & 2). Similar curves are obtained for the prokaryotic sample (not shown).

Interestingly, 35 out of the 36 erroneous predictions for the eukaryotic sequences fall on the N-terminal side of the correct cleavage site, mostly in the region -6 to -3 (30/36). About half of these result from matches with a higher score in the h-region but a lower one in the c-region than calculated for the correct site, whereas only 6 out of 36 have higher c- and lower h-region scores than the correct site. I have thus tried to improve the predictive accuracy in various ways, e.g. by multiplying the -3 and -1 weights or the whole c-region score by an extra factor, or by allowing a small variation in the distance between the h- and c-regions, but have not been able to obtain more than marginal improvements on the order of 2-4% in the overall success-rate.

The method described here not only allows prediction of the most likely cleavage site in new signal sequences, it also makes it possible to discriminate quite efficiently between putative signal sequences and the N-terminal regions of cytosolic proteins. The distribution of maximum scores for the eukaryotic signal sequences is shown in Fig.2, together with the

THIS PAGE BLANK (USPTO)

corresponding distribution obtained for a sample of 132 40-residues long N-terminal regions of cytosolic eukaryotic proteins (8). Only 3/161 (2%) of the signal sequences have maximum scores < 3.5; conversely, only 2/132 (2%) of the cytosolic sequences have maximum scores > 3.5. This level of discrimination compares favourably with that obtained with a recently published signal-sequence detecting algorithm (9).

DISCUSSION

Using a standard weight-matrix approach easily implemented even on a micro-computer, it is possible to set up a prediction method that (i) provides a clean discrimination between signal sequences and the N-terminal region in cytosolic proteins, and (ii) can be expected to identify the correct cleavage site 75-80% of the time when applied to new sequences not included in the data base (both prokaryotic and eukaryotic). This represents a significant improvement over previous methods.

Since the first submission of this work, another 36 eukaryotic signal sequences with known cleavage sites have been added to the data base. Using the same weight-matrix as above (Table 1), 75% of these sites were correctly predicted.

ACKNOWLEDGEMENT

This work was supported by a grant from the Swedish Natural Sciences Research Council.

REFERENCES

- (1) von Heijne, G. (1985) *J. Mol. Biol.* 184, 99-105.
- (2) von Heijne, G. (1983) *Eur. J. Biochem.* 133, 17-21.
- (3) Perlman, D., and Halvorson, H.O. (1983) *J. Mol. Biol.* 167, 391-409.
- (4) Mollay, C. (1985) In *The Enzymology of Post-translational Modification of Proteins*, Vol. 2, pp. 1-23, Academic Press, London.
- (5) Staden, R. (1984) *Nuc. Acids Res.* 12, 505-519.
- (6) Kuhn, A., & Wickner, W. (1985) *J. Biol. Chem.* 260, 15914-15918.
- (7) von Heijne, G. (1984) *J. Mol. Biol.* 173, 243-251.
- (8) Flinta, C., Persson, B., Jörnvall, H., and von Heijne, G. (1986) *Eur. J. Biochem.* 154, 193-196.
- (9) McGeoch, D.J. (1985) *Virus Res.* 3, 271-286.
- (10) Klapper, H.M. (1977) *Biochem. Biophys. Res. Commun.* 78, 1018-1024.

A calculation of fragment lengths obtainable from human DNA with 78 restriction enzymes: an aid for cloning and mapping

Radoje Drmanac, Nenad Petrović, Vladimir Glišin and Radomir Crkvenjakov

Genetic Engineering Center, PO Box 794, 11000 Belgrade, Yugoslavia

Submitted 13 May 1986

The optimal choice of restriction enzymes to be used with mammalian DNA is critical for construction of subgenomic libraries pulse field gel electrophoresis, RFLP studies and other applications. For this purpose we have grouped 78 restriction enzymes representing the majority of known recognition sequences according to the expected average fragment length S (in bp) and provided estimates of the number of different fragments N and the maximal length of fragments Dmax obtainable with each.

The frequencies of 16 possible dinucleotides in human DNA (Swartz, J., Trautner, M. and Kornberg, A. (1962) *J. Biol. Chem.*, 237, 1961.) do not meet the expectation of randomness for a DNA having a GC content of 39.4%. We have defined S as the inverse of the product of the frequencies of dinucleotides forming the restriction site and the probabilities of their association. Thus S is highly sequence dependent. The results are applicable to all mammalian DNAs, since their dinucleotide frequencies are similar. Due to the statistical nature of the calculation, the real values may differ if the sequences involved are not random for functional or structural reasons.

THIS PAGE BLANK (USPTO)

Mutations in Signal Sequence Cleavage Domain of Preproparathyroid Hormone Alter Protein Translocation, Signal Sequence Cleavage, and Membrane-Binding Properties

Kristine M. Wiren*, Lionel Ivashkiv, Philip Ma, Mason W. Freeman, John T. Potts, Jr., and Henry M. Kronenberg

Endocrine Unit
Massachusetts General Hospital
Department of Medicine
Harvard Medical School
Boston, Massachusetts 02114

Signal sequences, known to mediate the targeting of nascent secreted proteins to membranes, share common structural domains: a positively charged amino-terminus, a hydrophobic core, and a signal cleavage domain. Mutations have been introduced into the cDNA encoding the signal sequence of the mammalian protein preproparathyroid hormone to analyze the roles played by the signal cleavage domain in secretion. Two mutant genes were constructed, missing the entire six-residue propeptide sequence and several residues of the signal cleavage domain. The effects of these mutations on signal function were assessed after expression in clonal cell lines and in a transcription-linked translocation system. Alterations in the signal cleavage domain resulted in reduced translocation and signal cleavage. Furthermore, in one mutant the removal of the signal cleavage domain converted the signal into a membrane anchor sequence. The nonhydrophobic sequences at the end of the signal sequence thus crucially affect the translocation, cleavage, and membrane-binding properties of signal sequences. (Molecular Endocrinology 3: 240-250, 1989)

INTRODUCTION

All eukaryotic secretory proteins begin synthesis within the cytoplasm unattached to membrane. An amino-terminal peptide extension, termed the signal peptide, directs the nascent chain to the membrane of the rough endoplasmic reticulum (1, 2). The functional importance of the signal sequence has been established through

the analysis of mutant proteins. Characterization of many mutant bacterial proteins that fail to be secreted has shown that these proteins contain mutations in the signal sequence; further, fusion of a signal sequence from the precursor of a secreted eukaryotic protein to the sequence of a normally cytoplasmic protein caused the cytoplasmic protein to cross the rough endoplasmic reticular membrane (3). While the functional importance of the signal sequence has thus clearly been demonstrated, the specific role(s) played by the signal in directing proteins into the secretory pathway is (are) still not fully established.

The early portion of the secretory pathway can be broken conceptually into a series of processes. Signal sequences first bind to a signal recognition particle (SRP) (4), which, in turn, binds to an SRP receptor (docking protein) (5) in the membrane of the endoplasmic reticulum. Several models have been proposed for the subsequent insertion-translocation process. The nascent protein may bind initially to specific transport proteins (6); alternatively signal sequences may interact directly with the lipid bilayer (7), or may alter the structure of the nascent chain to achieve a conformation energetically competent for transport through the membrane (8). Tight binding of nascent proteins to membranes requires GTP (9, 10). The subsequent protein translocation is accompanied often by cleavage of the signal sequence by a membrane-bound enzyme, signal peptidase (11). If the signal sequence is not cleaved, it may pass completely through the membrane (12), or may instead anchor the protein in the membrane (13). Signals therefore may play several roles in the discrete steps in secretion, including binding to SRPs, insertion into the membrane, interaction with the translocation apparatus to effect translocation of attached sequences as well as of the signal itself, and cleavage by signal peptidase.

THIS PAGE BLANK (USPTO)

Comparisons of the sequences of many signal peptides have revealed no conserved primary amino acid sequence, but have revealed striking structural similarities among the signals that suggest the functional importance of the conserved domains. Three distinct domains have been noted (14, 15): a usually positively charged amino-terminal region, a central hydrophobic core of 10–15 amino acid residues, and a non- α -helical signal peptidase cleavage domain.

The functional importance of the three domains in the signal sequence has been demonstrated by characterization of mutant proteins. The positive charge at the amino-terminus is important for translocation of lipoprotein through the *Escherichia coli* membrane (16). Charged residues or α -helix breaking sequences in the hydrophobic core also affect transport of prokaryotic proteins (3). Alterations near the cleavage site can disrupt signal peptidase cleavage (17–20).

The putatively non- α -helical domain, which might introduce flexibility into the molecule, may allow signal peptides to adopt a loop or hairpin structure near the signal cleavage site in the membrane. The hairpin configuration has been proposed as the structure appropriate for insertion of the precursor into the membrane (13, 21–23) and for presenting an appropriate substrate to the signal peptidase (24–27). In order to analyze the possible roles played by the signal cleavage domain of a eukaryotic signal sequence, we have altered the sequence of cDNA encoding preproparathyroid hormone (preproPTH) in this region. PreproPTH, the precursor of PTH, contains a typical 25-residue amino-terminal signal sequence, followed by a six-residue prospecific peptide and the 84-residue mature hormone. PreproPTH cDNA has been cloned and expressed in clonal cell lines and in cell-free extracts in the presence of microsomal membranes. In the present study, we have constructed two mutant genes missing the entire propeptide sequence and the last several residues of the signal sequence. We have characterized the phenotypic consequences of these mutations and show that removal of residues from the signal cleavage domain can affect the efficiency of protein translocation and cleavage. Further, the uncleaved signal sequence of one mutant was converted into a membrane anchor sequence. The relatively subtle change required for this conversion suggests that the change in the predicted secondary structure of this mutant may play an important role in the phenotypic conversion.

RESULTS

Construction of Mutant PTH Precursor DNAs

The nucleotide and amino acid sequences of the wild type preproPTH cDNA and the mutant cDNAs employed in this study are illustrated in Fig. 1A. To create plasmid pPTH Δ L1, pPTHm124 (28) was cleaved at codons encoding the juncture between the prospecific sequence and the mature PTH sequence (Fig. 1B). The

free ends were progressively removed with Bal 31 exonuclease, and the shortened signal sequence was ligated with a synthetic oligonucleotide linker to the intact mature PTH sequence. The resulting mutant encodes a protein missing the last five residues of the signal sequence (which includes one residue from the hydrophobic core and the complete cleavage domain), and the entire pro-specific peptide. The introduced residues pro-gly, encoded by the synthetic linker, join the signal sequence (in phase) to the mature PTH sequence. Construction of the second mutant exploited convenient restriction endonuclease cleavage sites in the bovine preproPTH signal sequence (highly homologous to the human signal sequence) and at the end of the propeptide sequence in human preproPTH (Fig. 1B). After cleavage and separation of the fragments, the fragment encoding the first 22 residues of the bovine signal sequence was ligated directly to that encoding the mature human PTH sequence. The resulting mutant, called pPTH Δ L2, therefore contains an intact hydrophobic core, but is missing the last three residues of the signal sequence, as well as the pro-specific peptide. The sequence of both mutant PTH cDNAs were confirmed by DNA sequencing using the chemical method (29).

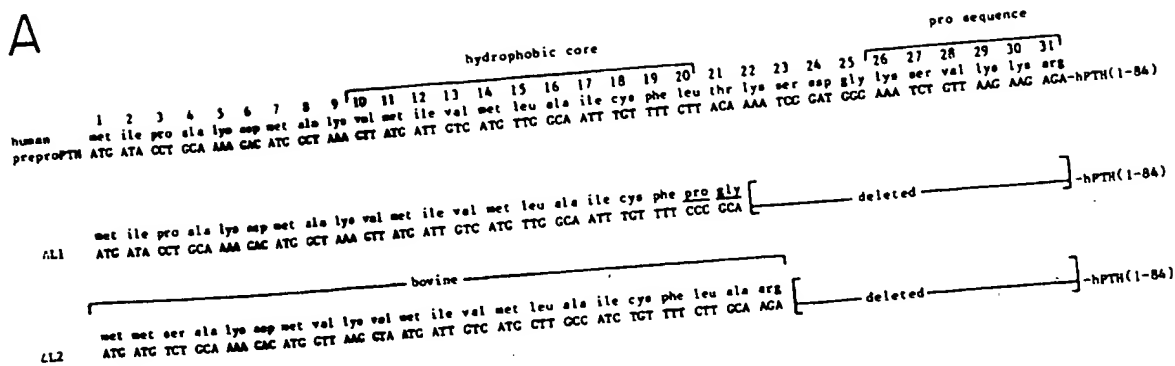
The mutant PTH cDNAs were then inserted into two expression vectors: a retroviral vector (pZIP-v-gpt) for studies in intact cells (a construct analogous to pZIP-neoSV(X)1 (30), but containing the gpt gene), and an *E. coli* expression plasmid pGL101 (31) for studies employing cell-free transcription-linked translation assays. These constructs are illustrated in Fig. 1C. GH₄C₁ cell lines stably expressing mutant Δ L1 or Δ L2 cDNA were obtained by infection of cultures with recombinant retrovirus containing the mutant cDNA, generated by transfection of ψ -2 cells (32). In the pGL101-based constructs, transcription of the mutant PTH mRNA was directed by the lac promoter *in vitro* using *E. coli* RNA polymerase. After adjusting buffer conditions (33), the mRNA was translated in rabbit reticulocyte lysate.

Processing of Mutant Precursors in GH₄C₁ Clonal Lines

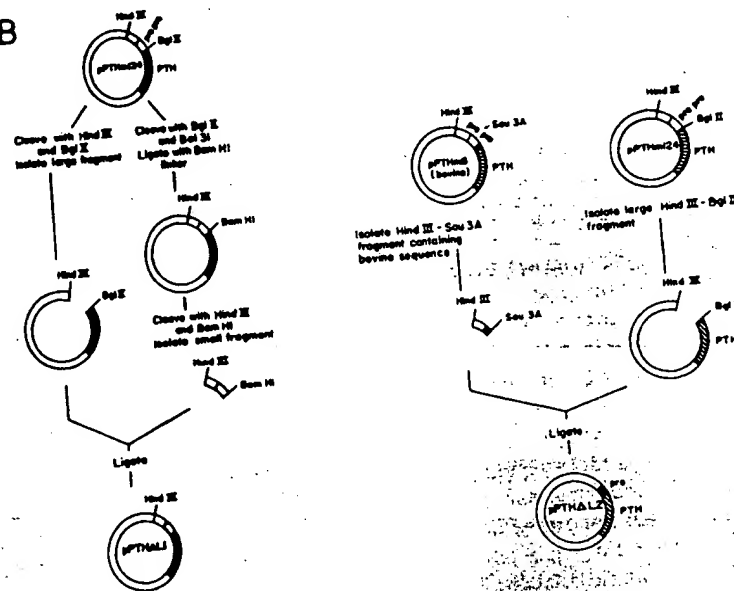
The two deletion mutants of preproPTH were analyzed for defects in processing and secretion in intact GH₄C₁ cells. Pulse-chase analyses of cell lines expressing preproPTH, the Δ L1 precursor, and the Δ L2 precursor were performed. Cells were pulse-labeled with [³⁵S] methionine for 15 min and chased for various times as listed in Fig. 2; PTH-related peptides were immunoprecipitated from both cell lysates and media, and visualized by fluorography after sodium dodecyl sulfate (SDS)-polyacrylamide electrophoresis. In the Δ L2-transformed cell line, the predominant protein immunoprecipitated from cell extracts was the size expected for the uncleaved precursor (Fig. 2A, lane 2) based on comigration with pre Δ L2 marker protein expressed in reticulocyte lysate (lane 1). In striking contrast, labeling of cells expressing the normal preproPTH yielded no

THIS PAGE BLANK (USPTO)

A



B



C

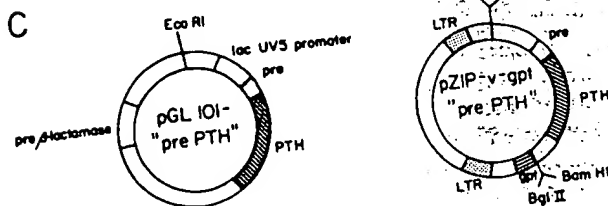


Fig. 1. Structures of Mutant Genes and Plasmid Constructions

A, Amino acid and nucleotide sequences of precursor-specific sequences of human preproPTH, preΔL1, and preΔL2. The underlined pro-gly residues of ΔL1 are encoded by a synthetic oligonucleotide. Each precursor-specific region abuts the identical mature human PTH sequence. B, Construction of ΔL1 plasmid (left) and of ΔL2 plasmid (right). C, Expression vectors for cell-free synthesis (left) and for production of recombinant retrovirus (right). Figures are not drawn to scale. To construct the retroviral plasmids, the Hind III sites of pΔL1 and pΔL2 were converted to Bam HI sites and the resultant Bam HI-ended fragments were inserted into the Bgl II site of pZIP-v-gpt.

intact precursor at the earliest time point, because the signal sequence was efficiently cleaved (34). Remarkably, the ΔL2 precursor was quite stable intracellularly, little changed after 3 h of chase (lane 6.) Radiosequence analysis of newly synthesized cellular protein isolated from a gel confirmed that this peptide had the labeled methionine residues at positions 1, 2, 7, 11, and 14

expected for the uncleaved preΔL2 precursor (data not shown). A small amount of the ΔL2 cleaved product was variably detectable in the cell extracts. This PTH-sized band was secreted into the media from cells more slowly than from cells synthesizing normal preproPTH (34), and somewhat more slowly than was the ΔL1 cleaved product (Fig. 2B). Amino-terminal radiosequ-

THIS PAGE BLANK (USPTO)

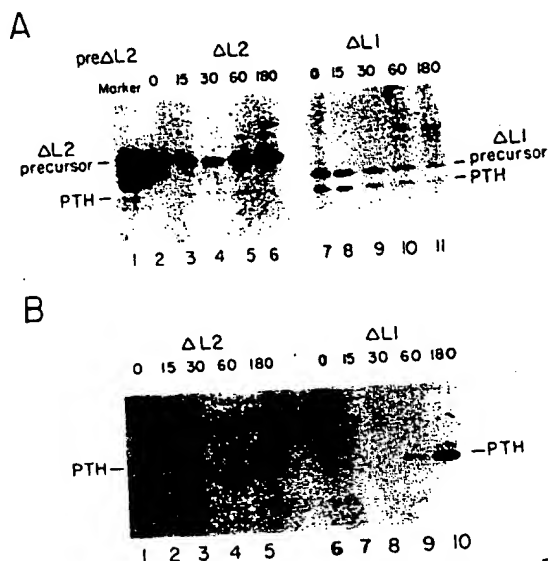


Fig. 2. Pulse-Chase Analysis of Clonal GH₄C₁ Cell Lines Expressing Δ L1 and Δ L2 Recombinant Retroviral Genes

Cells were pulsed with [35 S]methionine for 15 min and chased with nonradioactive methionine for the indicated times (in minutes). Cell lysates (A) and media (B) were immunoprecipitated and subjected to gel electrophoresis and fluorography.

ence analysis (Fig. 3C) demonstrated that the secreted protein in Δ L2-expressing cell lines was a mixture of two PTH-related peptides: PTH(2-84) and PTH(4-84); that is, PTH peptides beginning at either the second or fourth amino acid of the mature PTH molecule (Fig. 3D). Consequently, the methionines normally at positions 8 and 18 of PTH were found at positions 5 and 15, and 7 and 17 in the mixture of secreted peptides (Fig. 3C). Identical peaks were reproduced when this experiment was repeated (data not shown). No uncleaved precursor was secreted into the media from the cells (Fig. 2B).

After labeling for 15 min, the predominant immunoprecipitated peptide found in cells expressing the Δ L1 gene was the size expected for the uncleaved precursor (Fig. 2A, lane 7). Again, amino-terminal sequence analysis confirmed that this protein was the uncleaved precursor (data not shown). A substantially greater amount of a cleaved PTH-sized product found in the extract from Δ L1-expressing cells than in Δ L2-expressing cells (Fig. 2A). Sequence analysis showed that this peptide was PTH(4-84) (Fig. 3B). Thus, in intact cells, the pre Δ L1 precursor was cleaved more efficiently than the pre Δ L2 precursor but less efficiently than normal preproPTH. Moreover, the cleaved product of the pre Δ L1 precursor disappeared from cell extracts and appeared in the medium with more rapid kinetics. Unlike normal preproPTH, but like the pre Δ L2 precursor, the pre Δ L1 precursor was detected intracellularly. The stability of the Δ L1 precursor differed from that of the Δ L2 precursor, however, since the Δ L1 precursor was degraded rapidly intracellularly (Fig. 2A). The contrast between the stabilities of these similar precursor pep-

tides suggests the possibility that the two precursors might be in different intracellular compartments; the unstable pre Δ L1 precursor might be accessible to cytoplasmic proteases, while the stable pre Δ L2 precursor might be sequestered from the cytoplasm.

Processing of pre Δ L1 and pre Δ L2 Peptides in Cell-Free Extracts

A cell-free, transcription-linked translation system was used in order to more precisely localize products and characterize defects in the processing these mutants. Using the *lac* promoter-based expression vectors described in Fig. 1C, *E. coli* RNA polymerase transcribed mRNA for both the PTH-related genes and the pre- β -lactamase gene, the precursor of the antibiotic resistance protein. Pre- β -lactamase is transported across canine pancreatic microsomal membranes and accurately cleaved by signal peptidase (33-35). Consequently, monitoring the processing of the rough microsomes (RM), provided an internal control for the analysis of the processing mutant PTH precursors.

To assay processing by signal peptidase, proteins were translated from mRNA in the presence of [35 S]methionine, with or without RM, immunoprecipitated, separated on 15%-20% SDS-polyacrylamide gradient slab gels and processed for fluorography. Figure 4 compares the processing of normal human preproPTH, of the pre Δ L1 precursor, and of the pre Δ L2 precursor. In the absence of microsomes, the predominant immunoprecipitated protein from translation reactions programmed by the preproPTH-expression plasmid was preproPTH (lane 11). Other translation products immunoprecipitated by PTH antisera, migrating faster than preproPTH, resulted from initiation at the internal methionines at position 7 of the preproPTH signal sequence and at positions 8 and 18 of the mature PTH molecule (data not shown). The two signal-containing precursors in this translation reaction were converted to proPTH in the presence of RM (lane 12). Because proPTH contains only two methionines, while human preproPTH contains six methionines available for incorporation of [35 S]methionine (see Fig. 1), the conversion of preproPTH to proPTH was more efficient than simple comparison of band intensities alone suggests.

In the absence of RM, Δ L2 DNA (lane 6) directed the synthesis of the full-length precursor (sequence data shown in Fig. 3A), as well as the precursor beginning with methionine 7 (sequence data not shown). Lane 1 shows the analogous pre Δ L1 precursor synthesized in the absence of RM. Cleavage of both mutant precursors was considerably less efficient than the cleavage of normal preproPTH (compare lanes 2 and 7 to lane 12). In this particular experiment, the pre Δ L2 precursors were cleaved so poorly that cleaved product was just detectable above the PTH(8-84) band. The relative inefficiency of signal cleavage *in vitro* of the two mutant precursors correlates well with the results from intact cells: preproPTH > pre Δ L1 > pre Δ L2.

THIS PAGE BLANK (USPTO)

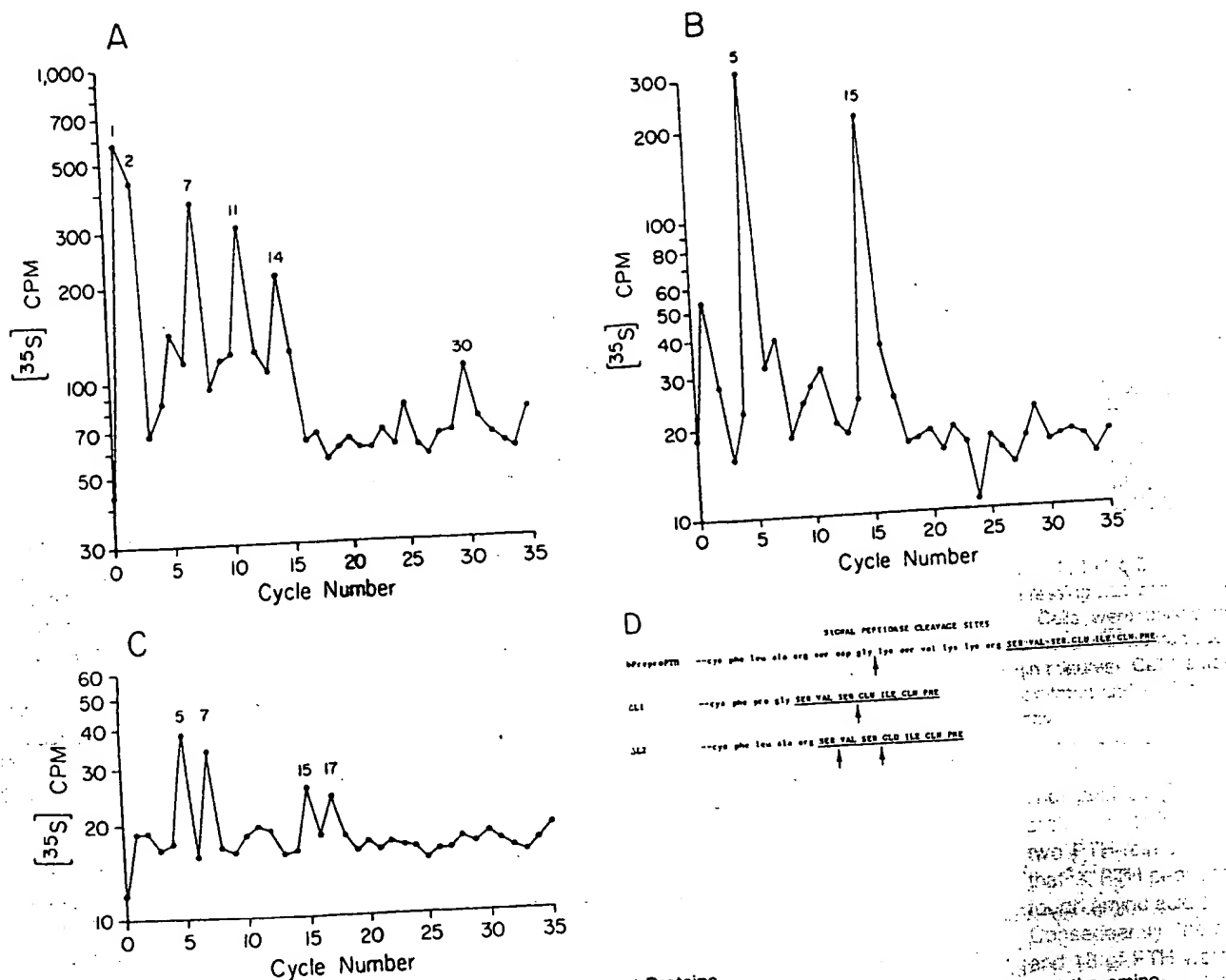


Fig. 3. Amino-Terminal Radiosequence Analysis of Expressed Mutant Proteins

A, Pre $\Delta L2$ precursor synthesized in cell-free extract. Methionines at 1, 2, 7, 11, and 14 correspond to those at the amino-terminus of the predicted sequence. B, PTH-sized protein secreted into media from cells expressing $\Delta L1$. Peaks at 5 and 15 demonstrate that the sequenced protein is PTH (4-84). C, PTH-sized protein secreted into media from cells expressing $\Delta L2$. Peaks at 5, 15, and 17 correspond to methionines normally at positions 8 and 18 of PTH (1-84). Consequently, the sequenced proteins represent a mixture of PTH (2-84) and PTH (4-84). The identical peaks were identified in a second, independent experiment. D, Cleavage sites of bovine preproPTH, pre $\Delta L2$ precursor, and pre $\Delta L1$ precursor.

Translocation of the Mutant Precursors through the M membrane

The observed inefficient cleavage of the mutant precursors could occur because the proteins were translocated inefficiently to the vicinity of the signal peptidase, or because the mutant sequences were poor substrates for signal peptidase. To distinguish between these possibilities, a proteolytic-protection assay was used to assess the ability of the mutant precursors to traverse the microsomal membrane. In this assay, proteases added to completed translation reactions digest proteins that have not been translocated across the microsomal membrane but cannot digest translocated proteins. After translation of protein in the presence of microsomal membranes, trypsin and chymotrypsin (100 $\mu\text{g/ml}$ each) were added for 90 min at 0 C. Products

remaining after proteolysis were immunoprecipitated, separated by SDS-polyacrylamide gel electrophoresis, and visualized by fluorography. To assure that the proteins remaining after proteolytic digestion were not inherently resistant to proteolysis, one control reaction contained both protease and 1% Triton X-100; sequestered proteins become sensitive to proteolysis after disruption of membrane integrity. As a further control, membrane and then proteases were added to translation reactions after translation was complete. In this way any nonspecific hydrophobic interaction that would result in interference with proteolysis in the presence of membranes could be detected.

Figure 4 shows the results of such a proteolytic-protection assay. ProPTH, cleaved from normal human preproPTH, was protected from digestion within the vesicular lumen, while precursor preproPTH was com-

THIS PAGE BLANK (USPTO)

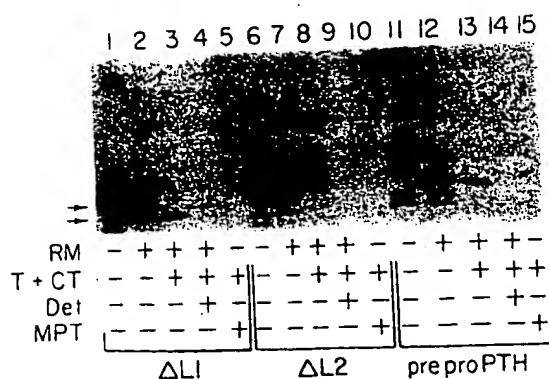


Fig. 4. Cleavage and Translocation of Related PTH Precursors
In indicated lanes RM were added during translation or posttranslationally (MPT). Trypsin and chymotrypsin were added to indicated lanes (T + CT); Triton X-100 was added along with protease as indicated (Det). Precursor and processed products are indicated by arrows.

pletely digested (lane 13). Similarly, processed $\Delta L1$ was protected, while the unprocessed precursor pre $\Delta L1$ was digested (lane 3). The inefficient cleavage of the pre $\Delta L1$ precursor was thus associated with inefficient translocation of the precursor through the membrane.

The translation products of the $\Delta L2$ expression plasmid exhibited a different pattern of protection from controlled proteolysis. Both of the pre $\Delta L2$ precursors were somewhat protected from proteolysis, as was the very small amount of cleaved protein, visible on longer exposure of the film (lane 8). These proteins were inherently digestible, since addition of detergent led to the proteolysis of the pre $\Delta L2$ precursors (lane 9). Furthermore, the protected precursors were not resistant as the result of a nonspecific hydrophobic interaction of the uncleaved signal with the membrane surface, since addition of RM after translation did not protect the pre $\Delta L2$ precursor from digestion (lane 10). Thus, although the pre $\Delta L2$ precursor was translocated through the membrane somewhat less efficiently than preproPTH (compare lanes 7 and 8 with lanes 12 and 13), the translocated protein was so inefficiently cleaved that the uncleaved, translocated precursor could be detected by the proteolytic-protection assay.

Location of the Precursor in the Membrane

Proteolytic-protection experiments are useful for distinguishing proteins that are outside from those that are inside the membrane-bound vesicle. However, they do not distinguish proteins that are inserted into the membrane from those that are translocated completely through the membrane and are free in the microsomal lumen or loosely bound to the membrane's inner surface. To determine the nature of the association of the translocated precursor with the membrane, the membranes were extracted under alkaline conditions and then centrifuged. Both loosely bound membrane proteins and those that are free within the lumen are extractable under such conditions (36). Only integral membrane proteins remain in the pelleted membran

fraction. After translation in the presence of RM and [35 S]methionine, reactions were adjusted to pH 11.5 with 1 M NaOH and incubated 10 min at 0°C (37). Membrane pellet and supernatant fractions were obtained by centrifugation, neutralized, immunoprecipitated, and subjected to SDS-polyacrylamide gel electrophoresis for analysis. Figure 5 demonstrates that, even under these vigorous extraction procedures, much of the precursor pre $\Delta L2$ translated in the presence of RM, was retained in the membrane fraction. Lane 8 contains the expected products of pre $\Delta L2$ translation in the presence of microsomes. The intensities of the bands in lane 8 are considerably lower than the intensities in lane 7, because a large amount of pancreatic microsomal extract was added to the translation reaction. The large amount was required because the cleavage and translocation of pre $\Delta L2$ are inefficient. These large amounts of membrane often nonspecifically inhibit protein synthesis. After centrifugation of the same amount of radioactive protein shown in lane 8, lanes 9 and 10 show that a substantial fraction of the pre $\Delta L2$ precursor pelleted with the membranes. This result suggests that at least a portion of the pre $\Delta L2$ precursor is an integral membrane protein. The association of the pre $\Delta L2$ precursor with the membrane occurred only when the membranes were present during translation. When membranes were added after translation and then extracted (lanes 11 and 12, which contain the same amount of radioactive protein as lane 7), only a small amount of precursor pre $\Delta L2$ associated nonspecifically with the membrane fraction and pelleted with the membrane (lane 12). Unlike pre $\Delta L2$, both preproPTH (data not shown) and pre $\Delta L1$ precursor proteins (lanes 3 and 4) were not associated with the membrane. The entire experiment in Fig. 5 was repeated a second time, with virtually identical results. Furthermore, experiments using differential urea extraction (38) to extract membrane proteins similarly demonstrated that pre $\Delta L2$ was tightly

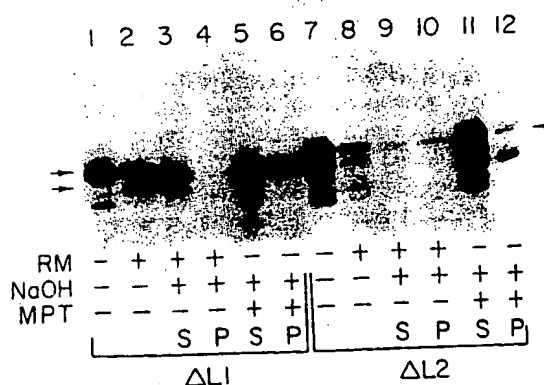


Fig. 5. Membrane Binding of Mutant PTH Precursors

RM were present during translation in indicated lanes or were added posttranslationally (MPT). Sodium hydroxide was added after translation to adjust pH to 11.5 in indicated lanes (NaOH). After addition of base and incubation on ice for 10 min, reactions were centrifuged. Supernatants (S) are displayed in lanes 3, 5, 9, and 11; pellets (P) are displayed in lanes 4, 6, 10, and 12. Precursor and processed products are indicated by arrows.

THIS PAGE BLANK (USPTO)

associated with the membrane (data not shown). Therefore, removal of several amino acids in the cleavage domain converted pre Δ L2 into an integral membrane protein.

DISCUSSION

Signal sequences have been characterized as containing three structural domains (14) that may have functional importance. In order to analyze the roles played by the hydrophilic signal cleavage domain in the secretory process, mutations were introduced into the cDNA of the eukaryotic secretory protein preproPTH. Two mutant genes were analyzed in the present study, both missing the entire propeptide-specific domain, and varying amounts of the signal cleavage domain.

The properties of the pre Δ L2 precursor (Fig. 1) differed from those of normal preproPTH in three respects: the precursor was translocated across the membrane somewhat inefficiently, the precursor was cleaved by signal peptidase with striking inefficiency and at novel residues, and much of the precursor became tightly associated with the microsomal membrane, with binding properties characteristic of integral membrane proteins. The pre Δ L1 precursor (Fig. 1) also exhibited an altered phenotype: the precursor was translocated across the membrane inefficiently; consequently, less than normal amounts of precursor were cleaved by signal peptidase and, again, cleavage was at a novel location (Fig. 3d). In contrast to the pre Δ L2 precursor, none of the pre Δ L1 precursor crossed the membrane without cleavage, and none could be shown to be tightly bound to the membrane. These results are consistent with earlier reports which demonstrated that mutations in the signal cleavage domain can disrupt cleavage of yeast (19, 20) and bacterial (39) precursor proteins.

The inefficient cleavage of the pre Δ L2 precursor is a particularly surprising observation. The sequence arg-ser-val-ser (Fig. 3D) adjacent to the hydrophobic core, provides a sequence that should well serve as an appropriate cleavage domain containing the small amino acids found at positions -1 and -3 adjacent to most cleavage sites (40). Nonetheless, while cleavage of the pre Δ L2 precursor did occur, it was strikingly inefficient. Application of the method of von Heijne (41) to predict the location of signal cleavage sites by comparing a test sequence with the sequences of a large number of known signal sequences predicts that, in fact, both cleavage sites used in the pre Δ L2 molecule should be well suited cleavage sites (Fig. 6). Since cleavage was so poor, we speculate that the sequence, though it has an appropriate primary sequence, did not, in fact, reach the signal peptidase in a suitable secondary structure. Perhaps the absence of a sufficiently flexible β -turn (see below) shielded the cleavage site from the enzyme. Lipp and Dobberstein (42) reported the behavior of a somewhat analogous sequence. They found that a cryptic signal sequence cleavage domain

could be made manifest by altering adjacent sequences of a membrane anchor sequence. They speculated that, in its normal setting, the cryptic site is not used because it does not reach the signal peptidase.

Analysis of the predicted secondary structure of the mutant and normal preproPTH molecule supports the possibility that the pre Δ L2 precursor might lack a suitably flexible signal sequence. Figure 7 displays the results of the application of the rules of Chou and Fasman (43) to the pre Δ L2 precursor (Fig. 7B), to human preproPTH (Fig. 7A), and to the pre Δ L1 precursor (Fig. 7C), using a computer program devised by Novotny and Auffray. The most striking difference observed in comparing the three sequences is in the predicted structure of the pre Δ L2 precursor: the pre Δ L2 precursor is missing a predicted β -turn structure that interrupts the α -helical region at the end of the signal sequence (predicted β -turns are represented by peaks above the line marked T.) The otherwise closely related pre Δ L1 precursor has β -turn potential, contributed by the pro-gly residues encoded by the synthetic oligonucleotide used in its construction. Similarly, normal preproPTH has the residues ser-asp-gly to provide β -turn potential at the end of the signal sequence. In the absence of a sequence which might introduce structural flexibility, the pre Δ L2 precursor might form a relatively uninterrupted extended α -helix, that could include the hydrophobic stretch from residues five to eight of mature PTH (ile-gly-leu-met).

The pre Δ L1 precursor was cleaved more efficiently contrantranslationally (in the presence of pancreatic microsomes), even though its primary structure resembles that of the pre Δ L2 precursor. Presumably, the flexibility

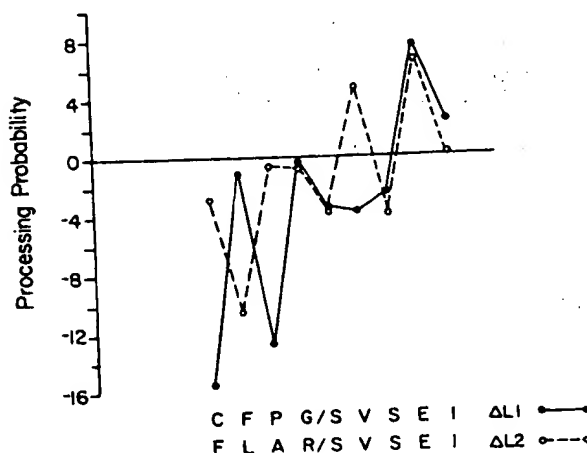


Fig. 6. Probability that Indicated Residues Represent Signal Cleavage Sites

The probability matrix of von Heijne (41) was used to assign cleavage site probabilities to the residues in the expected cleavage domains of pre Δ L1 and pre Δ L2. Amino acid residues are indicated by single letter code, starting with residue 19 of the signal sequence of pre Δ L2. Processing probability values are assigned to a given residue, with signal peptidase cleavage occurring between that residue and the one preceding it. *Slash lines* indicate borders of the signal sequence and mature PTH (1-84).

THIS PAGE BLANK (USPTO)

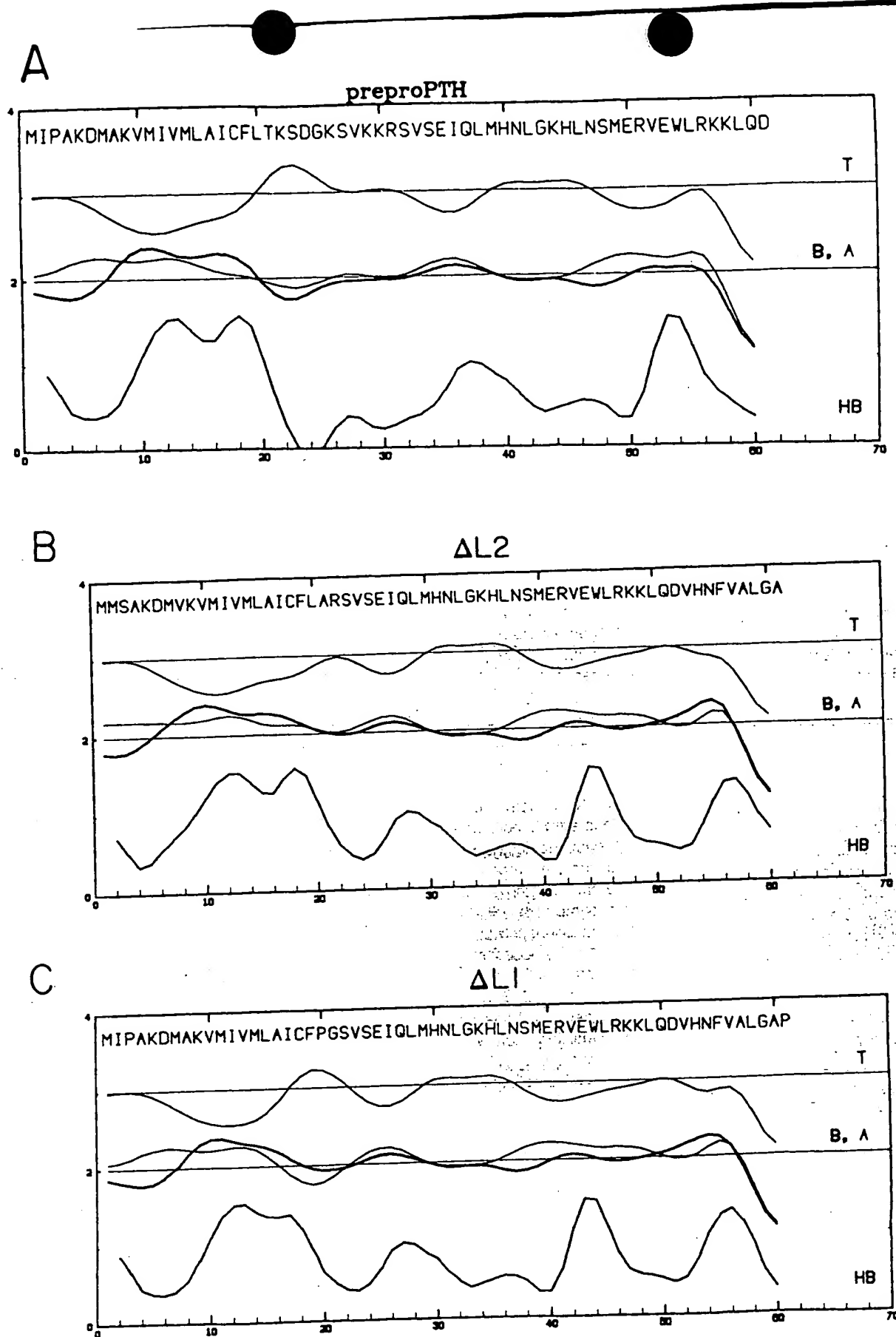


Fig. 7. Structural Predictions (51) for preproPTH, pre Δ L1, and pre Δ L2

Across the top are displayed the amino-terminal sequences of each precursor protein in single-letter code. The next line (T) tracks β -turn potential according to the predictions of Chou and Fasman (43); peaks above base-line have β -turn potential. The next line tracks both β -sheet potential (B, heavy line) and α -helical potential (A, thin line). Bottom curve tracks hydrophobic index (HB) as calculated by method based on Rose and Roy (43a).

THIS PAGE BLANK (USPTO)

provided by the inserted pro-gly residues (see Fig. 7) allowed cleavage after the second serine in the sequence pro-gly-ser-val-ser. The greater distance of the serine from the hydrophobic core may have encouraged cleavage as well. Interestingly, the pre Δ L1 precursor, unlike the pre Δ L2 precursor, was not cleaved after the first serine in this sequence. von Heijne (41) noted that proline is absent at position -3 from the signal cleavage sites of his large collection. As expected, this site has an unfavorable value in the calculated probability of processing using von Heijne's rules (41) (Fig. 6). The results observed with the pre Δ L1 mutant are therefore consistent with von Heijne's predictions for signal peptidase enzyme activities. Therefore, the observed differences in the cleavage sites chosen in the two mutants by signal peptidase may be explained by predicted signal peptidase specificities, while difference in cleavage efficiency may be due to alterations in the secondary structure of the mutant precursors.

The substantial inefficiency of cleavage of the pre Δ L2 precursor permits an analysis of the localization of the translocated but uncleaved precursor. Resistance of the precursor to proteolytic digestion and lack of extraction of the precursor from membranes with alkali suggest that the pre Δ L2 precursor was imbedded in the membrane as an integral membrane protein. The pre Δ L2 precursor's 12 residue hydrophobic core (shorter than a typical membrane anchor sequence) is separated from other hydrophobic residues in the mature PTH molecule by arginine and glutamic acid residues, as well as by two polar serine residues. Adams and Rose (44) showed that shortening of the membrane-spanning domain of the G protein of the vesicular stomatitis virus to twelve or even eight residues still allowed the protein to span the membrane. Davis and Model (45) have demonstrated that a hydrophobic stretch of as few as 10-14 residues can serve to anchor a precursor in the membrane, although unstably; the hydrophobic stretch in the signal of pre Δ L2 could well perform a similar function. Further, the extended α -helix predicted by the Chou-Fasman calculations might allow the hydrophobic residues at positions 5-8 of mature PTH to further stabilize the membrane-spanning domain of pre Δ L2.

The pre Δ L2 precursor, and more strikingly the pre Δ L1 precursor, were translocated across the microsomal vesicle less efficiently than was normal preproPTH (Fig. 4). The particularly inefficient translocation of the pre Δ L1 precursor may have been caused by the shortening of that precursor's hydrophobic core from 12 to 10 residues. (This calculation of core length assume that the proline introduced by the oligonucleotide into the pre Δ L1 sequence terminates the core.) In contrast, the sequence of the hydrophobic core of the pre Δ L2 precursor is normal; only the signal cleavage domain is altered. The somewhat inefficient translocation of this precursor may have been due to deletion of the propeptide-specific domain. Removal of this domain alone also resulted in somewhat inefficient translocation (Wiren, K. M., H. M. Kronenberg, and J. T. Potts Jr., submitted for publication). Alternatively, the lack of

predicted flexibility associated with the loss of the ser-asp-gly residues at the end of the signal sequence may have disrupted the protein's normal interaction with the translocation apparatus. It is also possible that the inability of signal peptidase to remove the signal sequence efficiently led to inefficient translocation of the rest of the protein.

We have found that alterations in the signal cleavage domain result in both reduced translocation and signal cleavage efficiency. Furthermore, in the Δ L2 mutant, the removal of a few amino acids in the signal cleavage domain converted the signal sequence into a membrane anchor sequence. In order to more fully understand the mechanisms responsible for the altered phenotypes of the mutant precursors, the several steps in the translocation and cleavage process must be assessed independently using appropriate *in vitro* test systems. Such an analysis of the mutants described here, and of other analogous structures, should lead to a greater understanding of the functions of discrete portions of signal sequences, and the relationships between signal sequences and stop-transfer sequences.

MATERIALS AND METHODS

Plasmid Construction

The methods for preparation of plasmid DNA, cleavage with restriction enzymes, purification of DNA fragments, ligation with T4 ligase and transfection of *E. coli* were performed as described (46). The Δ L1 mutant plasmid created from pPTHm124 was obtained after Bal 31 exonuclease digestion; the sequences of both mutants were determined by DNA sequence analysis (29). Plasmid pZIP-v-gpt was a gift of R. Mulligan (Department of Biology, MIT, Cambridge MA; and Whitehead Institute for Biomedical Research).

Cells

The rat pituitary cell line GH₄C₁ (47), was maintained in Dulbecco's modified Eagle's medium containing 10% calf serum in 95% air-5% CO₂. The ψ -2 packaging cell line (32), a clonal NIH 3T3 cell line productively transfected with a deleted Moloney murine leukemia virus missing packaging sequences, was a gift of R. Mulligan.

DNA Transfection and Viruses

DNA transfections were performed with 10 μ g plasmid DNA by the procedure of Graham and Van der Eb (48), as modified by Parker and Stark (49). Virus infections in the presence of 8 μ g Polybrene/ml for 2.5 h. Aminopterin was omitted from gpt-selection medium prepared for use with GH₄C₁ cells (50).

Protein Labeling and Sequence Determination

To radiolabel cells, 100-mm plates of confluent cells were rinsed in methionine-deficient DMEM containing 10% dialyzed fetal calf serum, and then incubated with 0.5 mCi [³⁵S]methionine in 4 ml of the same medium. Chase medium contained 30 mg nonradioactive methionine/liter. Incubations were terminated by cell lysis with detergent-containing buffer as described previously (34), and the protease inhibitor phenylmethylsulfonylfluoride (500 μ g/ml).

THIS PAGE BLANK (USPTO)

Immunoprecipitation with Rabbit Anti-Bovine PTH Antibody

Recovery of the immunoprecipitated proteins was accomplished as previously described (34). Proteins were separated on a 15%–20% SDS-polyacrylamide gradient slab gels. Gels were fixed, incubated in En³Hance (DuPont-NEN, Wilmington, DE) according to manufacturer's specifications, dried, and subjected to fluorography. Sequencing of proteins eluted from dried gels was by automated Edman degradation as previously described (33).

Transcription-Linked Translation

Expression plasmids, based on pGL101 (31) were transcribed by *E. coli* RNA polymerase as previously described (33). The mRNA produced was translated in a nuclease-digested reticulocyte lysate. To test for translocation through the membrane (of added RM from canine pancreas), limited protease digestion was done. Trypsin and chymotrypsin were added to the translation reactions at 100 µg/ml. Triton X-100, at 1% final concentration, was added as a control to some reactions. As an additional control for some reactions, membranes were added after translation was complete at 30 min. Reactions were then incubated 90 min on ice, followed by the addition of aprotinin (150 kallikrein-inhibiting units) to all reactions. Triton X-100 was then added at 0.2% final concentration to those tubes without prior detergent addition. Radiolabeled proteins thus produced were then subjected to immunoprecipitation as described above.

Membrane Association

Alkaline extraction of protein from membranes was based on the method of Russel and Model (37). After transcription and translation in the presence of membranes as described above, reactions were adjusted to pH 11.5 with 1 M NaOH and incubated on ice for 10 min. Membrane and supernatant fractions were obtained by centrifugation at 15,000 × g for 30 min. After immunoprecipitation, radiolabeled proteins were separated on 15%–20% SDS-polyacrylamide gradient slab gels. Gels were incubated in En³Hance and dried, and fluorography was performed.

Acknowledgments

We are grateful to Drs. J. Hendy and J. O'Riordan for the PTH antisera provided. We thank Dr. J. Novotny for this helpful discussion and aid in preparing Fig. 7. The authors are grateful for the excellent secretarial support provided by Susan Rhodes.

Received September 2, 1988. Revision received October 12, 1988. Accepted October 14, 1988.

Address requests for reprints to: Dr. Kristine M. Wren (current address), Geriatric Research, Education and Clinical Center, American Lake Veteran's Medical Center, Tacoma, Washington 98493.

This work was supported by NIH Grant AM-27758.

* Supported by National Research Service Award DK-07708 and NIH Training Grant AM-07028.

REFERENCES

1. Blobel G 1980 Intracellular protein topogenesis. *Proc Natl Acad Sci USA* 77:1496–1500
2. Wickner WT, Lodish HF 1985 Multiple mechanisms of protein insertion into and across membranes. *Science* 230:400–407
3. Lingappa VR, Chaidez J, Yost CS, Hedgpeth J 1984 Determinants for protein localization: β -lactamase signal sequence directs globin across microsomal membranes. *Proc Natl Acad Sci USA* 81:456–460
4. Walter P, Blobel G 1981 Translocation of proteins across the endoplasmic reticulum. III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *J Cell Biol* 91:557–61
5. Meyer DI, Krause E, Dobberstein B 1982 Secretory protein translocation across membranes—the role of the "docking protein." *Nature* 297:647–650
6. Wiedmann M, Kurzchalia TV, Hartmann E, Rapoport TA 1987 A signal sequence receptor in the endoplasmic reticulum membrane. *Nature* 328:830–833
7. Briggs MS, Cornell DG, Dluhy RA, Rierach LM 1986 Conformations of signal peptides induced by lipids suggest initial steps in protein export. *Science* 233:206–208
8. Wickner W 1979 The assembly of proteins into biological membranes: the membrane trigger hypothesis. *Annu Rev Biochem* 48:23–45
9. Mueckler M, Lodish HF 1986 Post-translational insertion of a fragment of the glucose transporter into microsomes requires phosphoanhydride bond cleavage. *Nature* 322:549–552
10. Connolly T, Gilmore R 1986 Formation of a functional ribosome-membrane junction during translocation requires the participation of a GTP-binding protein. *J Cell Biol* 103:2253–2261
11. Evans EA, Gilmore R, Blobel G 1986 Purification of microsomal signal peptidase as a complex. *Proc Natl Acad Sci USA* 83:581–585
12. Hortin G, Boime I 1981 Transport of an uncleaved preproteins into the endoplasmic reticulum of rat pituitary cells. *J Biol Chem* 256:1491–1494
13. Spiess M, Lodish HF 1986 An internal signal sequence: the asialoglycoprotein receptor membrane anchor. *Cell* 44:177–185
14. von Heijne G 1985 Signal sequences. The limits of variation. *J Mol Biol* 184: 99–105
15. Perlman D, Halvorson HO 1983 A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *J Mol Biol* 167:391–409
16. Vlasuk GP, Inouye S, Ito H, Itakura K, Inouye M 1983 Effects of the complete removal of the basic amino acid residues from the signal peptide on the secretion of lipoprotein in *Escherichia coli*. *J Biol Chem* 258:7141–7148
17. Inouye S, Duffaud G, Inouye M 1986 Structural requirement at the cleavage site for efficient processing of the lipoprotein secretory precursor of *Escherichia coli*. *J Biol Chem* 261:10970–10975
18. Koshland D, Sauer RT, Botstein D 1982 Diverse effects of mutations in the signal sequence on the secretion of β -lactamase in *Salmonella typhimurium*. *Cell* 30:903–914
19. Schauer I, Emr S, Gross C, Schekman R 1985 Invertase signal and mature sequence substitutions that delay intercompartmental transport of active enzyme. *J Cell Biol* 100:1664–1675
20. Hagenauer-Tsapis R, Hinnen A 1984 A deletion that includes the signal peptidase cleavage site impairs processing, glycosylation, and secretion of cell surface yeast acid phosphatase. *Mol Cell Biol* 4:2668–2675
21. Steiner DF, Quinn PS, Chan SJ, Marsh J, Tager HS 1980 Processing mechanisms in the biosynthesis of proteins. *Ann NY Acad Sci* 343:1–16
22. Sabatini DD, Kreibich G, Morimoto T, Adesnik M 1982 Mechanisms of the incorporation of proteins in membranes and organelles. *J Cell Biol* 92:1–22
23. Engelman DM, Steitz TA 1981 The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. *Cell* 23:411–422

THIS PAGE BLANK (USPTO)

24. Folz RJ, Gordon JI 1986 Deletion of the propeptide from human preproapolipoprotein A-II redirects contranlational processing by signal peptidase. *J Biol Chem* 261:14752-14759
25. Finidori J, Rizzolo L, Gonzalez A, Kreibich G, Adesnik M, Sabatini DD 1987 The influenza hemagglutinin insertion signal is not cleaved and does not halt translocation when presented to the endoplasmic reticulum membrane as part of a translocating polypeptide. *J Cell Biol* 104:1705-1714
26. Coleman J, Inukai M, Inouye M 1985 Dual functions of the signal peptide in protein transfer across the membrane. *Cell* 43:351-360
27. Inouye S, Duffaud G, Inouye M 1986 Structural requirement at the cleavage site for efficient processing of the lipoprotein secretory precursor of *Escherichia coli*. *J Biol Chem* 261:10970-10975
28. Born W, Freeman M, Hendy GN, Rapoport A, Rich A, Potts Jr JT, Kronenberg HM 1987 Human preproparathyroid hormone synthesized in *Escherichia coli* is transported to the surface of the bacterial inner membrane but not processed to the mature hormone. *Mol Endocrinol* 1:5-14
29. Maxam A, Gilbert W 1980 Sequencing end-labelled DNA with base-specific chemical cleavages. *Methods Enzymol* 65:499-560
30. Cepko CL, Roberts BE, Mulligan RC 1984 Construction and applications of a highly transmissible murine retrovirus shuttle vector. *Cell* 37:1053-1062
31. Guarente L, Lauer G, Roberts TM, Ptashne M 1980 Improved methods for maximizing expression of a cloned gene: a bacterium that synthesizes rabbit β -globin. *Cell* 20:543-553
32. Marin R, Mulligan RC, Baltimore D 1983 Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* 33:153-159
33. Kronenberg HM, Fennick BJ, Vasicsek TJ 1983 Transport and cleavage of bacterial pre- β -lactamase by mammalian microsomes. *J Cell Biol* 96:1117-1119
34. Helleman JG, Cone RC, Potts Jr JT, Rich A, Mulligan RC, Kronenberg HM 1984 Secretion of human parathyroid hormone from rat pituitary cells infected with a recombinant retrovirus encoding preproparathyroid hormone. *Proc Natl Acad Sci USA* 81:5340-5344
35. Mueller M, Ibrahim I, Chang CN, Walter P, Blobel G 1982 A bacterial secretory protein requires signal recognition particle for translocation across mammalian endoplasmic reticulum. *J Biol Chem* 257:11860-11863
36. Steck TL, Yu J 1973 Selective solubilization of proteins from red blood cell membranes by protein perturbants. *J Supramol Struct* 1:220-248
37. Russel M, Model P 1982 Filamentous phage pre-coat is an integral membrane protein: analysis by a new method of membrane preparation. *Cell* 28:177-184
38. Gilmore R, Blobel G 1985 Translocation of secretory proteins across the microsomal membrane occurs through an environment accessible to aqueous perturbants. *Cell* 42:497-505
39. Kuhn A, Wickner W 1985 Conserved residues of the leader peptide are essential for cleavage by leader peptidase. *J Biol Chem* 260:15914-14918
40. von Heijne G 1983 Patterns of amino acids near signal-sequence cleavage sites. *Eur J Biochem* 133:17-21
41. von Heijne G 1986 A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res* 14:4683-4690
42. Lipp J, Dobberstein B 1986 The membrane-spanning segment of invariant chain (1) contains a potentially cleavable signal sequence. *Cell* 46:1103-1112
43. Chou PY, Fasman GD 1978 Empirical predictions of protein conformation. *Annu Rev Biochem* 47:251-276
- 43a. Rose GD, Roy S 1980 Hydrophobic basis of packing in globular proteins. *Proc Natl Acad Sci USA* 77:4643-4647
44. Adams GA, Rose JK 1985 Structural requirements of a membrane-spanning domain for protein anchoring and cell surface transport. *Cell* 41:1007-1015
45. Davis NG, Model P 1985 An artificial anchor domain: hydrophobicity suffices to stop transfer. *Cell* 41:607-614
46. Hendy GN, Kronenberg HK, Potts Jr JT, Rich A 1981 Nucleotide sequence of cloned cDNAs encoding human preproparathyroid hormone. *Proc Natl Acad Sci USA* 78:7365-7369
47. Tashjian AH, Lomedico ME, Maina D 1978 Role of calcium in the thyrotropin releasing hormone-stimulated release of prolactin from pituitary cells in culture. *Biochem Biophys Res Commun* 81:798-806
48. Graham R, Van der Eb A 1973 A new technique for the assay of infectivity of human adenovirus DNA. *Virology* 31:360-369
49. Parker BA, Stark GR 1979 Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by the virus or viral DNA. *J Virology* 52:456-457
50. Mulligan RC, Berg P 1981 Selection for animal cells that express the *Escherichia coli* gene coding xanthine-guanosine phosphoribosyltransferase. *Proc Natl Acad Sci USA* 78:2072-2076
51. Novonty J, Auffray C 1984 A program for prediction of protein secondary structure from nucleotide sequence data: application to histocompatibility antigens. *Nucleic Acids Res* 12:243-255



THIS PAGE BLANK (USPTO)

Substrate Specificity of Eukaryotic Signal Peptidase

SITE-SATURATION MUTAGENESIS AT POSITION -1 REGULATES CLEAVAGE BETWEEN MULTIPLE SITES IN HUMAN PRE(Δ PRO)APOLIPOPROTEIN A-II*

(Received for publication, August 12, 1987)

Rodney J. Folz†§, Steven F. Nothwehr†, and Jeffrey I. Gordon†||

From the Departments of †Biological Chemistry and ‡Medicine, Washington University School of Medicine, St. Louis, Missouri 63110

We have previously described a novel mutant of human preproapolipoprotein A-II (pre(Δ pro)apoA-II) in which the wild-type 18-amino acid-long signal sequence (Gly¹⁸¹) was functionally redefined to 20 amino acids in length (Ala²⁰¹). We have used this mutant as a model preprotein to probe the substrate specificity of eukaryotic signal peptidase. Site-saturation mutagenesis was performed resulting in the substitution of 13 different amino acids (acidic, basic, aromatic, hydrophobic, and small-neutral) for Ala²⁰ (or position -1). The effects of these substitutions were assessed using an *in vitro* transcription/translation/microsomal membrane processing system. NH₂-terminal sequence analysis of the 13 mutant proteins demonstrated that amino acids which occupy position -1 in a signal peptide are critical in establishing a good context for signal peptidase cleavage and that two or more potential sites for cleavage may compete for recognition by signal peptidase.

Eukaryotic signal peptidase is an integral membrane protein complex of the endoplasmic reticulum (Kreibich *et al.*, 1980; Lively and Walsh, 1983) capable of endoproteolytically (Strauss *et al.*, 1979; Mollay *et al.*, 1982; Stern and Jackson, 1985) cleaving signal peptides from preproteins during co- and post-translational translocation across the endoplasmic reticulum bilayer (Blobel and Dobberstein, 1975; Caulfield *et al.*, 1986; Maher and Singer, 1986). Recently, canine pancreatic microsomal signal peptidase (Evans *et al.*, 1986) as well as hen oviduct signal peptidase (Baker *et al.*, 1986) have been purified. They exist as a complex of six polypeptides which require phospholipids for biological activity (Jackson and White, 1981). Proteolytic processing of prokaryotic signal peptides is mediated by two proteases: signal peptidase I (Wolfe *et al.*, 1983) and signal peptidase II (Innis *et al.*, 1984). Bacterial signal peptidase II is sensitive to the antibiotic globomycin (Dev *et al.*, 1985) and only cleaves preproteins which contain lipids coupled to their NH₂-terminal residues. Bacterial signal peptidase I is not affected by globomycin and cleaves the precursors of most other exported prokaryotic proteins (for review, see Wu, 1986). In addition, bacterial signal peptidase I correctly processes some eukaryotic preproteins (e.g. Watts *et al.*, 1983) as well as hybrid prepolypeptides

containing prokaryotic signal peptides fused to eukaryotic cytosolic proteins (Lingappa *et al.*, 1984). The converse has also been demonstrated for eukaryotic signal peptidase which correctly processes a bacterial preprotein that is normally cleaved by bacterial signal peptidase I (Muller *et al.*, 1982). Recently, Garcia *et al.* (1987) constructed several mutants of a hybrid protein containing the signal peptide and the first nine amino acids from bacterial outer membrane lipoprotein fused directly to β -lactamase. They examined the site of eukaryotic signal peptidase cleavage of this chimeric mutant and compared these results to those obtained in a prokaryotic system (Ghrayeb *et al.*, 1985). They concluded that eukaryotic signal peptidase and prokaryotic signal peptidase I possess similar substrate specificities which contrast markedly with the specificity of bacterial signal peptidase II (Garcia *et al.*, 1987).

Amino acid sequence comparisons of known eukaryotic signal peptides have provided clues about structural features which may be important in defining the substrate specificity of eukaryotic signal peptidase (Austen, 1979; Perlman and Halvorson, 1983; von Heijne, 1983). These studies suggest that several features of the primary and secondary structure of prepeptides may affect efficient and accurate cleavage of signal peptides by signal peptidase (review in von Heijne, 1984, 1985). For example, a β -turn structure 5-6 residues upstream of the site of signal peptidase cleavage has been predicted to occur in most eukaryotic signal sequences (Austen, 1979; Perlman and Halvorson, 1983). This β -turn interrupts the α -helical or β -strand structure of the hydrophobic membrane-spanning domain present in prepeptides and presumably allows, promotes, or presents a conformation compatible with signal peptide cleavage. The primary amino acid sequence of the signal peptide may also contribute to signal peptidase specificity: amino acids found at positions -1 and -3 almost always contain small-neutral side chains (e.g. Ala, Gly, Ser, Thr, and Cys), although position -3 is somewhat less restrictive and may include hydrophobic residues (e.g. Leu, Val, Ile). This observation has been formally incorporated into the "(-3,-1) rule" by von Heijne (1983, 1984) or the "A-X-B¹" model by Perlman and Halvorson (1983).

Inspection of eukaryotic preprotein data bases has yielded rules for predicting sites of co-translational cleavage. However, there have been no systematic attempts to determine experimentally the relative contribution various amino acids may have at different positions within the signal peptide on cleavage site selection by eukaryotic signal peptidase. Although the substrate requirements appear to be similar in eukaryotes and prokaryotes, more information is available about the prokaryotic signal peptidase. For example, Pollitt *et al.* (1986) have shown that substituting two amino acids of increasing size at position -1 in the bacterial lipoprotein

* This work was supported in part by National Institutes of Health Grant HL-18577 and by the Monsanto Corp. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported in part by Medical Scientist Training Program Grant GM-07200.

|| Established Investigator of the American Heart Association.

THIS PAGE BLANK (USPTO)

signal peptide prevents cleavage by signal peptidase II. Dierstein and Wickner (1986) demonstrated that residues spanning positions -9 to +5 of the M13 procoat protein influence the efficiency of signal peptidase I processing. Inouye *et al.* (1986) and Lehnhardt *et al.* (1987) have proposed a structural compatibility model for prokaryotic preprotein processing: for any given signal peptide to be fully functional, it must be structurally compatible with the protein to be secreted. They proposed that this structural compatibility is based on secondary structure present at or around the site of signal peptide cleavage.

In a recent study, we described a mutant eukaryotic preprotein that was constructed by deleting the prosegment coding region from a human preproapolioprotein A-II cDNA (Folz and Gordon, 1986). PreproapoA-II contains an 18-residue prepeptide, a pentapeptide prosegment, and a 77-residue mature domain which is incorporated into human high density lipoproteins. Although the 18-amino acid-long signal peptide was left completely intact in the mutant pre(Δ pro)protein, co-translational cleavage studies demonstrated that the site of cleavage was redirected to a unique position two amino acids into the mature protein domain (Gly¹⁸-Gln-Ala²⁰¹). Based on this observation, we anticipated that judiciously placed mutations near or around the site of signal peptide cleavage in pre(Δ pro)apoA-II would help elucidate structural requirements important in determining the site of preprotein processing. The strength of this model preprotein rests on the fact that if a mutation is made which inhibits cleavage at the primary site of signal peptidase processing (Ala²⁰¹), an alternative, secondary site of cleavage is potentially available (Gly¹⁸¹). The results of the mutagenesis can therefore be interpreted by comparing the extent of cleavage at positions 18, 20, or 18 and 20.

We have now substituted 13 different amino acids for Ala²⁰ (or position -1) in pre(Δ pro)apoA-II and studied the effects of these mutations on *in vitro* co-translational translocation and signal peptidase processing. Based on these studies, we demonstrate that amino acids at position -1 are critical in defining a good context for eukaryotic signal peptidase cleavage. Furthermore, we show that two or more potential sites of signal peptidase cleavage may compete for recognition by this protease. Pre(Δ pro)apoA-II and the various mutants described here may be used to probe selectively the contribution of other structural features have in defining the site of signal peptidase cleavage.

EXPERIMENTAL PROCEDURES

Materials and Radiochemicals—pGEM2 DNA, T₇ RNA polymerase, nuclease-treated rabbit reticulocyte lysate, RNasin, and reagents for dideoxynucleotide sequencing were from Promega Biotec. T₄ DNA ligase and T₄ polynucleotide kinase were purchased from Bethesda Research Laboratories and New England Biolabs, respectively. L-[3,4-³H]valine (45 Ci/mmol) was obtained from Amersham Corp. L-[³⁵S]Methionine (1130–1150 Ci/mmol) was from Du Pont-New England Nuclear. Canine pancreatic microsomal membranes were either from Amersham Corp. or prepared as previously described by Walter and Blobel (1983).

Oligodeoxynucleotide-directed Site-saturation Mutagenesis of Human Pre(Δ pro)apoA-II—The overall strategy for mutagenesis is outlined in Fig. 1. First, a *StyI* site was engineered near the signal peptide processing site of pre(Δ pro)apoA-II by annealing ~0.08 pmol of single-stranded M13mp18 DNA containing human pre(Δ pro)apoA-II cDNA (Folz and Gordon, 1986) to 10 pmol of a mutagenic 33-base oligodeoxynucleotide which contains a single base mismatch (5'-GCCTTGAAGGACAGGCCAAGGAGCCATGTGTGG-3'). Hybridization occurred in a solution of 100 mM NaCl, 10 mM Tris-HCl (pH 8.0), and 1 mM EDTA. This hybrid was then used directly to transform *Escherichia coli* strain JM109 (Hanahan, 1985) without any prior *in vitro* primer extension (Burke and Olson, 1986). Duplicate plaque lifts of the resulting recombinant phage library were prehy-

bridized for 2.5 h at 60 °C in buffer A (0.9 M NaCl, 90 mM sodium citrate, 1% (w/v) Sarkosyl, and 50 µg/ml salmon sperm DNA). Filters were then placed in fresh buffer A containing the ³²P-labeled 33-mer (specific activity ~2.5 × 10⁶ cpm/pmol, 5 × 10⁶ cpm/ml of hybridization solution). Hybridization was allowed to proceed for 20 h at 60 °C. Filters were repetitively washed in buffer B (150 mM NaCl, 15 mM sodium citrate, 1% Sarkosyl) at increasing temperatures (25–68 °C). DNA was prepared from one of the candidate (probe-positive) mutants, and its cDNA insert was characterized by digestion with *StyI* as well as by nucleotide sequence analysis (Sanger *et al.*, 1977).

An *EcoRI* fragment of the pre(Δ pro)apoA-II cDNA containing the *StyI* site was subcloned into the *EcoRI* restriction site of pGEM2. The *PstI* site present in the pGEM2 polylinker was destroyed by double digestion with *AccI* and *HindIII*. The resulting DNA was treated with DNA polymerase I (Klenow fragment) and all four deoxynucleotide triphosphates to create blunt ends. A blunt-ended ligation was performed with T₄ DNA ligase, and the DNA was redigested with *HindIII* to reduce the background of recircularized vector DNA before transformation of *E. coli* strain DH5. A recombinant plasmid lacking the pGEM2 polylinker *PstI* site was identified. The relative orientation of its insert DNA was determined by digestion with *PstI*.

Thirteen amino acids were substituted at position 20 of pre(Δ pro)apoA-II using a cassette mutagenesis strategy (see Fig. 2). Digestion of the plasmid containing pre(Δ pro)apoA-II cDNA with *PstI* and *StyI* produced a 24-base gap which includes codons 14–21. Complementary degenerate oligodeoxynucleotides of 16 bases (5'-GCCTTGAAGGACAGNN-3') and 24 bases (5'-CTTGNNCTGTCCTTCAAGGCTGCA-3') were synthesized (as described below) where N at the first and second positions of codon 20 represents A, G, C, and T. The oligodeoxynucleotides were end-labeled with T₄ polynucleotide kinase in the presence of [³²P]ATP. The duplex synthetic oligonucleotide was ligated into the gap generated by *PstI* and *StyI* digestion of the recombinant plasmid. This was followed by transformation of *E. coli* strain DH5 (Hanahan, 1985). These cassettes replaced the coding sequences removed by *PstI/StyI* digestion and generated all possible substitution mutations at the first and second positions of codon 20. Mutated DNAs were identified by double-stranded dideoxy-DNA sequencing using the T₇ promoter primer on minilysate preparations (Birnboim and Doly, 1979) of plasmid DNA.

To generate an Ala²⁰ → Gly²⁰ mutant of pre(Δ pro)apoA-II, the 24-base oligodeoxynucleotide (5'-CTTGCCCTGTCCTTCAAGGCTGCA-3') containing a Gly²⁰ codon was used in place of the degenerate 24-base oligodeoxynucleotide. All other procedures were identical to those described in the preceding paragraph.

All oligodeoxynucleotides were synthesized by an Applied Biosystems Model 380A or -B DNA synthesizer employing the solid-phase phosphoramidite method. Synthesis of degenerate oligodeoxynucleotides in which all 4 bases were required at a particular position was accomplished by programming the delivery of all bases or by manually preparing a mixture containing all 4 bases before synthesis. The 16-mer degenerate oligodeoxynucleotide listed above was synthesized on an oxidizable solid support resin (Molecular Biosystems Inc.). This oligodeoxynucleotide was chemically deconjugated from the support resin as follows. Fibers were removed following the completion of DNA synthesis, extensively washed in distilled water, and submerged in acetonitrile. They were then placed in a glass tube containing a 10-ml solution of sodium metaperiodate (110 mg) and 20 mM Na₂PO₄ (pH 7.0). The tubes were sealed, wrapped in aluminum foil to prevent exposure to light, and rotated at room temperature for 2.5 h. The fibers were removed and washed extensively in distilled water. The oligodeoxynucleotide was solubilized by twice incubating the fibers with a 1-ml solution of water:*N*-propylamine (9:1) for 2.5 h at 50 °C with occasional mixing. The water:*N*-propylamine was removed by lyophilization, and the oligodeoxynucleotide was purified by successive ethanol precipitation.

In Vitro Transcription Reactions—Wild-type human preproapoA-II mRNA and pre(Δ pro)apoA-II mRNA were transcribed by T₇ RNA polymerase. All 14 pre(Δ pro)apoA-II site-saturation mutants were linearized with the restriction enzyme *NheI*, and a run-off transcription was performed essentially as described (Folz and Gordon, 1986) except that T₇ RNA polymerase (0.5 unit/µl) was used in place of SP6 RNA polymerase. RNA integrity and homogeneity were confirmed by electrophoresis of the ³²P-labeled RNAs through 4% polyacrylamide (3.8% polyacrylamide, 0.2% bisacrylamide) gels containing 8 M urea.

In Vitro Translation and Co-translational Translocation/Processing

THIS PAGE BLANK (USPTO)

Assay—T₇-derived RNA was translated in an *in vitro* nuclease-treated rabbit reticulocyte lysate cell-free translation system in the presence of [³⁵S]methionine, as described in an earlier publication (Folz and Gordon, 1986). To assay the kinetics of co-translational processing, 30-μl reactions containing [³⁵S]methionine (37 μCi), RNasin (1.3 units/μl), canine pancreatic microsomal membranes (10 μl), nuclease-treated rabbit reticulocyte lysate (10 μl), and mRNA (30 fmol/μl) were incubated at room temperature. At 30, 60, and 90 min, 10-μl aliquots were removed and added to 1 ml of ice-cold immunoadsorption buffer (Gordon *et al.*, 1983b) to stop the reaction. Rabbit anti-human apoA-II sera preadsorbed to protein A-Sepharose CL-4B was used to immunopurify the translation products (Gordon *et al.*, 1983a). Antigen-antibody-protein A-Sepharose complexes were extensively washed (Gordon *et al.*, 1983b), dissociated, and electrophoresed through a 16% polyacrylamide gel containing 0.1% SDS¹ (Laemmli, 1970). The slab gels were fixed (in 40% methanol, 10% acetic acid), dried, and autoradiographed by using Kodak XAR-5 film and a Lightning Plus screen (Du Pont-New England Nuclear) at -70 °C.

NH₂-terminal Protein Sequencing—[³⁵S]Methionine- or [³H]valine-labeled apoA-II polypeptides synthesized in reticulocyte cell-free translation systems containing canine pancreatic microsomal membranes were purified by immunoprecipitation and electrophoresed through denaturing 15% polyacrylamide tube gels containing 0.1% SDS (Laemmli, 1970). The tube gels were frozen and sliced (into 1-mm sections), and the radiolabeled protein was passively eluted overnight into 5 mM sodium phosphate buffer (pH 7.1) containing 1 mM phenylmethylsulfonyl fluoride and 1 mM benzamidine HCl. The ³⁵S- and ³H-labeled polypeptides were subjected to automated sequential Edman degradation using a 0.33 M Quadrol program (Thomas *et al.*, 1981; Gordon *et al.*, 1982) and a Beckman 890C spinning cup sequencer.

Data Collection and Computer-assisted Analysis—[³⁵S]Methionine-labeled protein bands from autoradiograms were scanned using a LKB Ultrascan XL laser densitometer, and the peaks were quantitated by integration using the densitometer's internal digital integrator. Only signals in the linear range of film sensitivity were used in the subsequent analysis. The primary translation products of all mutant mRNAs contain 2 methionine residues, whereas the processed proteins (which have lost their signal peptide) contain only 1 methionine. Therefore, the percent processing for each precursor protein was calculated as follows: % processing = (2 × processed apoA-II) / ((2 × processed apoA-II) + (unprocessed preapoA-II)) × 100. The computer programs SIGSEQ1 and SIGSEQ2 (Folz and Gordon, 1987) were used to predict signal peptidase processing probabilities for all mutant proteins described in this report.

RESULTS

Construction and Characterization of Site-saturation Mutants—The plasmid pAII(ΔPro) contains a cDNA which encodes the mutant human apolipoprotein pre(Δpro)apoA-II (see Fig. 1 and Folz and Gordon, 1986). To understand better the structural features important in defining the site of signal peptidase cleavage in this model preprotein, we performed site-saturation mutagenesis at the Ala²⁰ codon. The strategy we chose for constructing 13 different amino acid substitutions for Ala²⁰ is depicted in Figs. 1 and 2. The creation of a silent *StyI* restriction site in pre(Δpro)apoA-II cDNA involved a single base substitution (A to C) at the third position of codon 20 using the mutagenesis protocol of Burke and Olson (1986). The advantage of this protocol is that fewer spurious or unwanted mutations occur relative to other methods (Burke and Olson, 1986). An *EcoRI* fragment of pre(Δpro)apoA-II cDNA containing this unique *StyI* restriction site was subcloned into the *EcoRI* site of pGEM2. Next, a *PstI* site present in the polylinker of the pGEM2 vector was removed by digestion with *HindIII* and *AccI*. The resulting DNA overhang was filled in with DNA polymerase I, followed by ligation with T₄ DNA ligase. Mutants lacking the *PstI* polylinker site were characterized by restriction enzyme analysis. A small DNA fragment was removed from one of these mutants by digestion with *PstI* and *StyI*, generating a

pre(ΔPro)apo AII cDNA

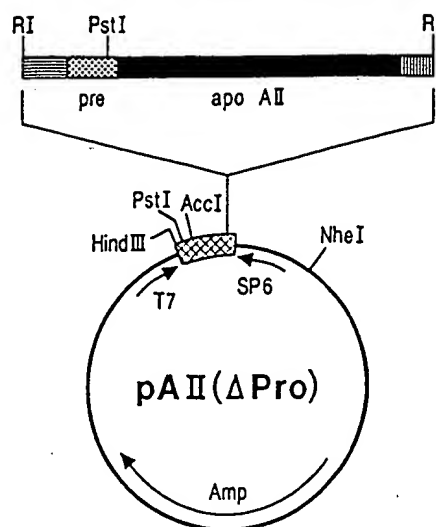


FIG. 1. Schematic drawing of the *in vitro* expression vector. Human pre(Δpro)apoA-II cDNA (which lacks a 15 base propeptide coding region (Folz and Gordon, 1986)) was subcloned into the *EcoRI* site of the pGEM2 expression vector to generate pAII(ΔPro). pAII(ΔPro) is the parent plasmid used for all subsequent mutagenesis. Nucleotide sequence regions include: ▨, 5'-nontranslated region; ▤, signal peptide (pre) coding segment; ■, mature or processed apoA-II coding segment; ▤, 3'-nontranslated region; and ▨, pGEM2 polylinker. The thin line represents pGEM2 plasmid sequence. *RI*, *EcoRI*; *Amp*, ampicillin resistance gene; *T*₇ and *SP6*, RNA polymerase promoter regions.

gap of 24 bases (Fig. 2). Oligodeoxynucleotides were synthesized that contained approximately equal amounts of all 4 bases at positions 1 and 2 of codon 20 as well as the adjacent codons (Fig. 2). The synthetic duplex DNA was ligated into the 24-base gap and used to transform *E. coli* strain DH5, thereby generating a library of substitution mutations at codon 20.

We isolated 30 random mutant colonies from this library and sequenced the inserts contained in their recombinant plasmid DNAs. Of the 30 mutants characterized in this fashion, 27 contained only substitution mutations at positions 1 and 2 of codon 20, whereas three contained unusual or spurious DNA insertions or deletions located at or near the cassette site. From the 27 mutations which were characterized, we identified all possible amino acid substitutions which could occur at codon 20 (Fig. 2), except those coding for Gly and Phe (Fig. 3). To generate a Gly²⁰ mutant, we synthesized an oligodeoxynucleotide which specifically encodes Gly at codon 20 and followed the same cassette mutagenesis procedure previously outlined. DNA sequence analysis of six potential Gly²⁰ mutants revealed that all six encoded Gly at this position (Fig. 3). We made no further attempts to identify a Phe²⁰ mutant.

In Vitro Co-translational Translocation Efficiency—Amino acid mutations, either in the signal peptide itself or in the mature protein domain, may affect the overall efficiency of co-translational translocation and processing. If the mutation blocks any productive association of the precursor protein with the translocation apparatus and translocation does not occur, then by definition signal peptide cleavage of this mutant precursor is blocked because the precursor protein is never presented to signal peptidase. In order to examine the effects of these mutations on the efficiency of co-translational translocation/processing, an *in vitro* transcription/transla-

¹ The abbreviation used is: SDS, sodium dodecyl sulfate.



THIS PAGE BLANK (USPTO)

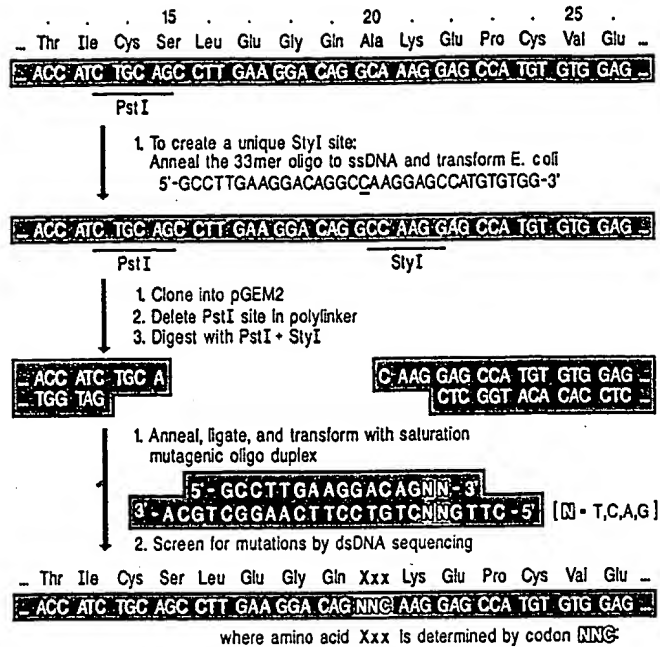


FIG. 2. Strategy for the construction of site-saturation mutants at codon 20 of human pre(Δ pro)apoA-II. A silent *Styl* restriction site was constructed in a pre(Δ pro)apoA-II cDNA by oligodeoxynucleotide-directed mutagenesis (for details, see "Experimental Procedures"). The cDNA was subsequently cloned into the *Eco*RI site of pGEM2. After deleting the *Pst*I site present in the pGEM2 polylinker, a DNA cassette was removed by *Styl* and *Pst*I digestion, resulting in a gap of 24 bases. Two complementary mutagenic oligodeoxynucleotides were synthesized such that positions 1 and 2 of codon 20 contained all 4 bases. This duplex oligodeoxynucleotide was ligated into the vector, followed by transformation into *E. coli* strain DH5 to generate a library containing 14 possible mutations at codon 20. Double-stranded DNA sequencing was performed to identify various mutants present in this library. All the possible amino acid substitution mutations shown in the box were isolated and characterized, except for Gly and Phe. The construction of a Gly²⁰ mutant was accomplished by the synthesis of a specific oligodeoxynucleotide (see "Experimental Procedures"). The numbering starts with the initiator methionine of pre(Δ pro)apoA-II.

tion/microsomal membrane processing system was employed (Folz and Gordon, 1986). T₇ RNA polymerase was used to generate homogeneous mRNA preparations encoding each of the 14 pre(Δ pro)apoA-II mutants as well as wild-type preproapoA-II. These 15 mRNAs were translated in parallel reticulocyte lysate cell-free systems containing [³⁵S]methionine and translocation-competent canine pancreatic microsomes. At 30, 60, and 90 min after initiation of translation, an aliquot from each cell-free reaction mixture was removed, and apoA-II was immunopurified and subjected to electrophoresis through 15% polyacrylamide gels containing 0.1% SDS (Laemmli, 1970). The processed and unprocessed apoA-II protein species were quantitated by laser densitometric scanning of autoradiographs of the slab gel. The percent processing was calculated for each precursor protein after correcting for the loss of the initiator methionine present in the signal peptide (see "Experimental Procedures"). Fig. 4 shows typical

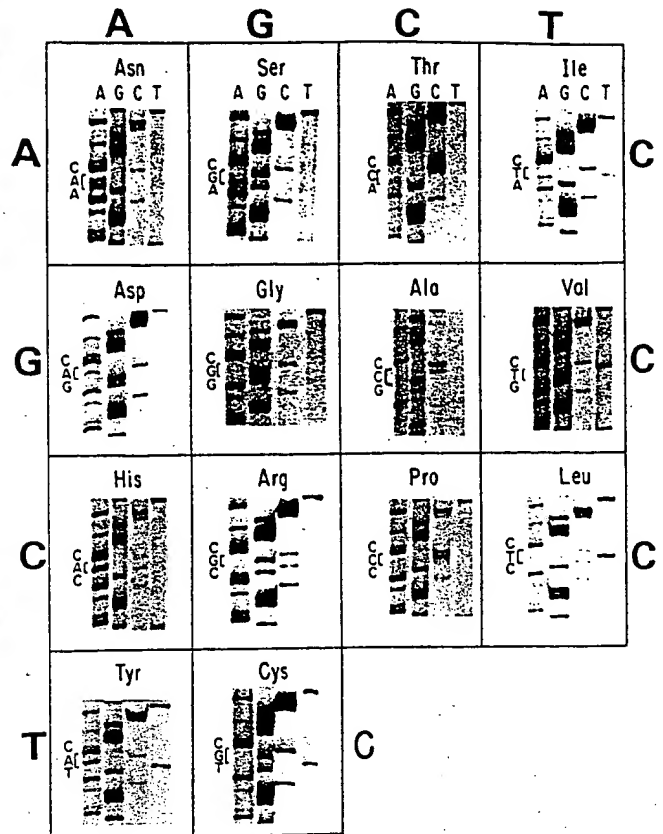


FIG. 3. Characterization of 14 pre(Δ pro)apoA-II site-saturation mutants at codon 20 by DNA nucleotide sequence analysis. The plasmid DNAs corresponding to each mutant shown were alkali-denatured, and the DNA sequence was determined using a T₇-promoter primer (5'-TATAGTGAGTCGTATTA-3'), DNA polymerase I (Klenow fragment), and the dideoxy chain termination method. Only relevant segments of the autoradiogram for each DNA sequence are shown. The deduced nucleotide sequences at the side of each gel represent (from the bottom to the top) nucleotides 1-3 of codon 20.

results obtained for wild-type preproapoA-II and four pre(Δ pro)apoA-II mutants; Ala²⁰, Pro²⁰, Tyr²⁰, and Arg²⁰. Wild-type preproapoA-II was the most efficiently processed precursor protein for all time points analyzed (Fig. 4 and data not shown). These four mutants were chosen because they are representative of the three groups of kinetic processing into which all 14 pre(Δ pro)apoA-II mutants could be classified (see Table I). Group I included those mutants which had the lowest percent processing at the 30-min time point relative to preproapoA-II. Members included Cys²⁰, Ser²⁰, Thr²⁰, Asn²⁰, Asp²⁰, Ile²⁰, and Tyr²⁰ pre(Δ pro)apoA-II. Group III (Ala²⁰, Pro²⁰, and Val²⁰) had the highest percent processing at 30 min, whereas Group II (Gly²⁰, His²⁰, Arg²⁰, and Leu²⁰) exhibited percent processing values that were intermediate between Groups I and III. As shown in Fig. 4, the percent processing values were highly dependent on the length of the incubation with the rate of processing increasing with time. Interestingly, Ala²⁰ is the only precursor mutation which shows no significant change in the rate of processing between 30 and 90 min.

Identification of the Site of Signal Peptidase Processing—As noted above, deletion of the propeptide from human preproapoA-II redirects signal peptide cleavage to a unique site two amino acids into the mature protein domain (after Ala²⁰¹) even though the 18-amino acid-long signal peptide was left completely intact. Thus, in this sequence context, cleavage

THIS PAGE BLANK (USPTO)

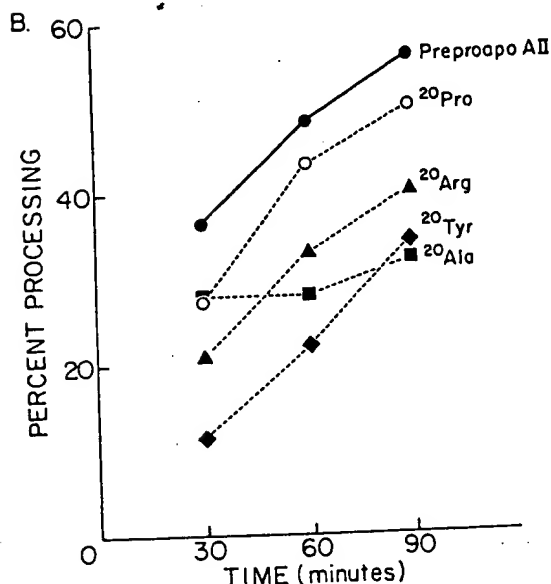
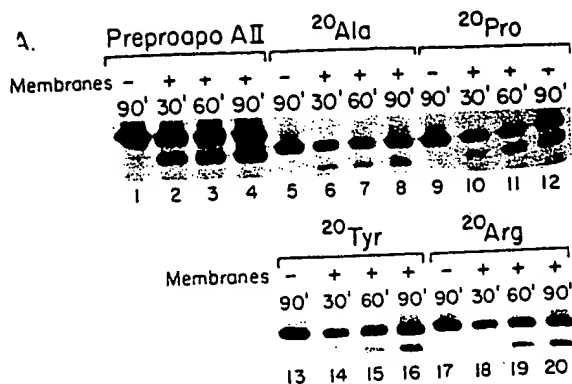


FIG. 4. Kinetics of co-translational signal peptidase processing. A, parallel reticulocyte lysate cell-free translation mixtures containing [35 S]methionine with or without canine pancreatic microsomal membranes (as indicated) were programmed with either wild-type preproapoA-II (lanes 1-4) mRNA or one of the following pre(Δ pro)apoA-II mutant mRNAs: Ala 20 (lanes 5-8), Pro 20 (lanes 9-12), Arg 20 (lanes 13-16), and Tyr 20 (lanes 17-20). Aliquots were removed at the time points indicated. The [35 S]methionine-labeled polypeptide samples were analyzed by denaturing SDS-polyacrylamide gel electrophoresis, followed by autoradiography. Bands corresponding to processed and unprocessed proteins were quantitated by laser densitometry for each time point indicated, and the percent processing was calculated (see "Experimental Procedures") and plotted as shown in B.

after Ala 201 must be favored over cleavage after the wild-type site (Gly 181). To determine if these two sites may compete for signal peptidase cleavage in different sequence contexts, we identified the site of signal peptidase cleavage for all 14 pre(Δ pro)apoA-II mutants by automated sequential Edman degradation. Fig. 5 illustrates potential sites of signal peptidase cleavage for the 14 pre(Δ pro)apoA-II precursor mutants. Inspection of the amino acid sequences indicates that if cleavage occurs after Xaa 201 , peaks of radioactivity should occur at cycles 5, 9, 16, and 27 for a [3 H]valine-labeled protein and at cycle 24 for a [35 S]methionine-labeled polypeptide. If signal peptidase cleavage occurs after Gly 181 , peaks of [3 H]valine would be expected at cycles 7, 11, and 18 and at cycle 26 for the [35 S]methionine-labeled protein. These distributions of radioactive peaks can be clearly distinguished from peaks arising from the unprocessed primary translation products which should yield [3 H]valine peaks at cycles 8 and 25 and [35 S]methionine at cycle 1.

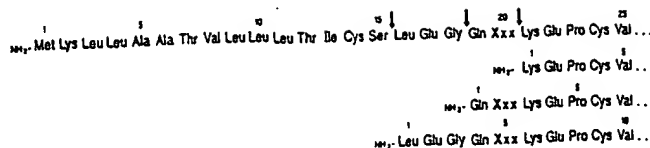


FIG. 5. Potential signal peptidase cleavage sites in human pre(Δ pro)apoA-II. The three most likely sites of predicted signal peptidase cleavage are indicated by the arrows (see Table I for a description of the method for predicting the site of signal peptidase cleavage). Xxx represents each of the 14 different amino acid substitutions found at codon 20.

With these possibilities in mind, *in vitro* transcribed mRNAs encoding all 14 pre(Δ pro)apoA-II mutants were each translated in reticulocyte lysate cell-free translation systems containing translocation-competent canine pancreatic microsomal membranes and [35 S]methionine or [3 H]valine. Immunoreactive processed and unprocessed apoA-II protein species were further purified by polyacrylamide gel electrophoresis (see "Experimental Procedures"). Fig. 6 displays representative NH $_2$ -terminal sequencing data obtained for each of the 14 mutants. For example, in Fig. 6A, the dominant peaks of [3 H]valine which occurred at cycles 5, 9, 16, and 27 and the single [35 S]methionine peak at cycle 24 could only have arisen if signal peptidase cleavage occurs after Cys 201 . Minor peaks at cycles 11 and 18 (not indicated in Fig. 6) which represent <5% of the total processed protein radioactivity suggest that less than 5% of signal peptidase cleavage occurred after Gly 181 in this Cys 20 mutant. In Fig. 6B, dominant peaks of [3 H]valine at cycles 2, 7, 11, and 18 and a [35 S]methionine peak at cycle 26 demonstrate that signal peptidase cleavage occurs only after Gly 181 in the mutant which has a valine at position 20. No other peaks of radioactivity corresponding to processed protein could be detected, suggesting that in this sequence context, signal peptidase cleavage after Val 201 is not very favorable (relative to cleavage after Gly 181). Since this Val 20 mutant substitutes a valine for alanine at codon 20 and since [3 H]valine is used to label the *in vitro* translation products, peaks of radioactivity due specifically to this mutation (as opposed to the other 13) occur at cycles 2 and 20 for the processed and unprocessed protein, respectively. In the Gly 20 mutant, signal peptidase cleavage occurs after both Gly 181 and Gly 201 (Fig. 6C). In this particular example, we calculated that 53% of the signal peptidase cleavage occurred after Gly 201 and 47% after Gly 181 . The data shown in Fig. 6D for the Leu 20 mutant provides some evidence for signal peptidase cleavage after Ser 181 (<5%, see Fig. 5), although the majority of cleavage (>95%) occurs after Gly 181 . There was no evidence for signal peptidase cleavage after Leu 201 .

A summary of the cleavage data obtained for all 14 pre(Δ pro)apoA-II mutants is provided in Table I. Recently, a new method for predicting the site of signal peptidase cleavage in preproteins has been developed by von Heijne (1986). These rules, which have been estimated to be about 80% accurate, are based on a statistical weight-matrix scoring system that was derived from an analysis of 161 known eukaryotic preproteins. The site with the highest score (see legend to Table I) is the site of predicted signal peptidase cleavage. We applied this set of rules to each of the 14 pre(Δ pro)apoA-II mutants. Comparison of the predicted *versus* observed site(s) of signal peptidase cleavage demonstrates some unexpected results (Table I). For example, 5-8% of the signal peptide cleavage occurred after Pro 201 . The remaining cleavage (92-95%) was after Gly 181 . Signal peptidase cleavage after a proline residue was originally thought to be prohibited (von Heijne, 1983);

THIS PAGE BLANK (USPTO)

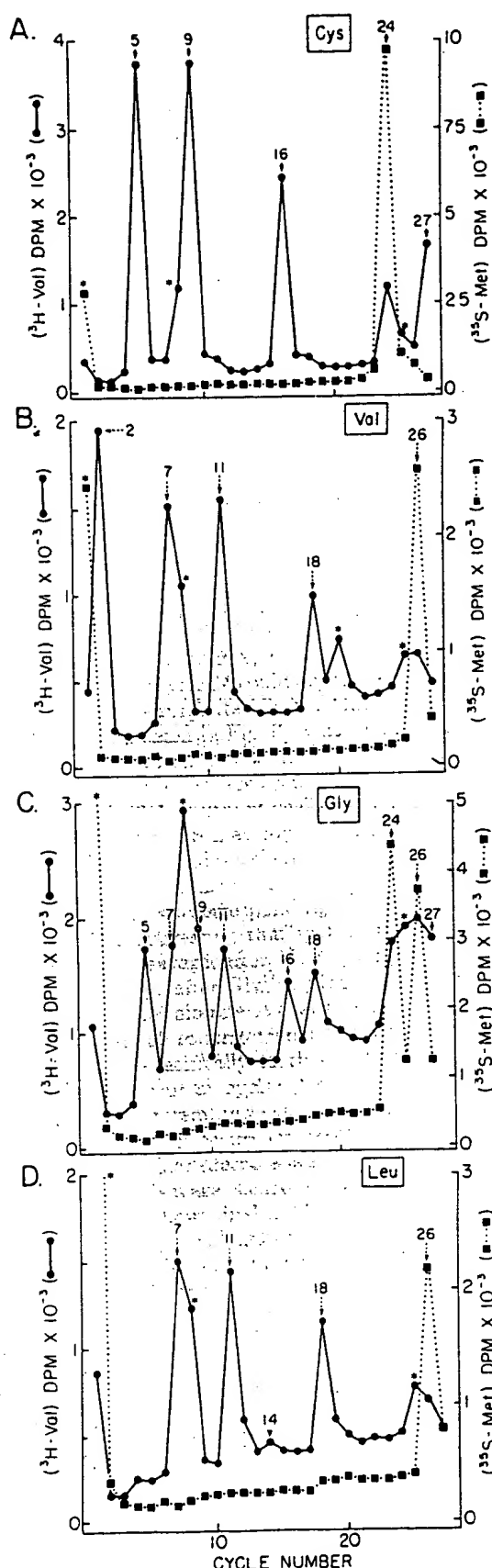


FIG. 6. NH₂-terminal protein sequence analysis of *in vitro* co-translationally cleaved pre(Δ pro)apoA-II mutants. *In vitro* transcribed mRNAs encoding the pre(Δ pro)apoA-II mutants were translated in separate reticulocyte lysate cell-free translation reaction mixtures supplemented with canine pancreatic microsomal membranes and [³⁵S]methionine or [³H]valine. Immunoreactive, radiola-

but the data presented here, as well as elsewhere (Weisman *et al.*, 1986; Pluckthun and Knowles, 1987), clearly indicate that a signal peptide may contain a proline at position -1 and still be a substrate for cleavage. Cleavage after Cys²⁰¹ was also not predicted to occur. Algorithms that compute the likelihood for preprotein processing at various sites are not able to distinguish whether processing occurs at a unique *versus* multiple positions (von Heijne, 1983, 1986). Therefore, they did not predict that the Cys²⁰, Gly²⁰, Ser²⁰, Thr²⁰, Pro²⁰, Ile²⁰, or Leu²⁰ mutants would be cleaved at two sites.

DISCUSSION

We have performed site-saturation mutagenesis at position -1 of a model preprotein: pre(Δ pro)apoA-II. The results indicate that altering the physical-chemical properties of residues located at this position in this sequence context has profound effects on signal peptidase cleavage specificity, with selection occurring at one or two sites for a given mutant.

Signal peptides probably perform multiple roles in protein export (Stader *et al.*, 1986). Kaiser *et al.* (1987) have shown that 20% of random polypeptide sequences can functionally replace the signal peptide of yeast preinvertase, suggesting that signal peptides are recognized by the export machinery through a mechanism that apparently has low specificity. Weisman *et al.* (1986) have indicated that the signal peptide of prelysozyme c can sustain one out of three randomly placed substitution mutations without loss of function. Despite these observations and despite the fact that there are no obvious consensus or canonical sequences for signal peptidase processing (Watson, 1984), examples of signal peptidase cleavage occurring at alternative or multiple sites that are different from the wild-type site are rare. In addition to pre(Δ pro)apoA-II, we are aware of only five other examples of eukaryotic signal peptidase cleavage at a site other than the wild-type site. A naturally occurring example of "ambiguous" cleavage is found with bovine pregrowth hormone, where 65% of the cleavage occurs after an Ala residue (Ala¹-Gly) and 35% takes place after an adjacent Gly (Ala-Gly¹) (Lingappa *et al.*, 1977). Hortin and Boime (1981) noted that incorporation of a threonine analog, β -hydroxynorvaline, into preprolactin synthesized by pituitary cells resulted in partial (~50%) redirection of signal peptide cleavage from the Thr located at position -1 to an Asn located three amino acids upstream (Asn¹-Val-Gln-Thr¹). They felt that this reflected steric constraints at the normal processing site imposed by the analog. Schauer *et al.* (1985) isolated a mutant (SUC2-s1) which contains an Ala to Val substitution at position -1 of the signal peptide of yeast preinvertase. Analysis of this mutant precursor protein demonstrated that the site of signal peptide cleavage had shifted one amino acid downstream of the wild-type site, occurring after an adjacent Ser (Val-Ser¹). Weisman *et al.* (1986) presented evidence which suggests that a mutation at position

beled apoA-II polypeptides were electrophoresed through 15% polyacrylamide tube gels containing SDS (0.1%). Specific bands of radioactivity corresponding to processed [³⁵S]methionine or [³H]valine protein products were eluted, pooled, and subjected to automated sequential Edman degradation. In A, 60,000 dpm of [³H]valine- and 60,000 dpm of [³⁵S]methionine-labeled Cys²⁰ apoA-II were sequenced. In B, 80,000 dpm of [³H]valine- and 60,000 dpm of [³⁵S]methionine-labeled Val²⁰ apoA-II were sequenced. In C, 98,600 dpm of [³⁵S]methionine-labeled Gly²⁰ apoA-II were sequenced. In D, 80,000 dpm of [³H]valine- and 80,000 dpm of [³⁵S]methionine-labeled Leu²⁰ apoA-II were sequenced. Peaks of radioactivity marked with a solid arrow demonstrate cleavage after Xaa²⁰¹. Dashed arrows indicate cleavage after Gly¹⁵¹. The solid arrowhead marks a peak of radioactivity consistent with cleavage after Ser¹⁵¹. An asterisk corresponds to peaks of radioactivity due to unprocessed precursor protein.

THIS PAGE BLANK (USPTO)

TABLE I
Human pre(Δ pro)apoA-II mutant phenotypes

Mutation ^a	Cleavage after ^b			Predicted site of cleavage ^c			Processing efficiency (% of control at 30 min) ^d
	Ser ¹⁵ ↓	Gly ¹⁸ ↓	Xaa ²⁰ ↓	Ser ¹⁵ ↓	Gly ¹⁸ ↓	Xaa ²⁰ ↓	
		%					
Ala ²⁰	0	0	100			↓	+++ (51)
Cys ²⁰	0	<5	>95		↓		+ (<5)
Gly ²⁰	0	47	53			↓	++ (32)
Ser ²⁰	0	74	26		↓		+ (<5)
Thr ²⁰	0	74	26		↓		+ (<5)
Pro ²⁰	0	92-95	5-8*		↓		+++ (37)
Asn ²⁰	0	100	0		↓		+ (<5)
His ²⁰	0	100	0		↓		++ (28)
Arg ²⁰	0	100	0		↓		++ (20)
Asp ²⁰	0	100	0		↓		+ (<5)
Ile ²⁰	<5	>95	0		↓		+ (<5)
Leu ²⁰	<5	>95	0	↓			++ (11)
Tyr ²⁰	0	100	0		↓		+ (<5)
Val ²⁰	0	100	0		↓		+++ (47)

^a In this report, individual mutants are identified as follows. The parent pre(Δ pro)apoA-II cDNA codes for an Ala at codon 20 (i.e. position -1 relative to the site of signal peptide cleavage in this mutant) and is referred to as Ala²⁰. A substitution of any of the 13 amino acids for Ala at codon 20 is referred to as Xaa²⁰, where Xaa represents the three-letter amino acid code for that particular mutant.

^b Based on NH₂-terminal sequencing of the [³⁵S]methionine- and [³H]valine-labeled processed proteins and taking into account a repetitive yield of 95% for each cycle of Edman degradation (see text and Fig. 6).

^c Predictions were based upon a set of rules described by von Heijne (1986). A statistical weight-matrix approach was used to calculate processing probability values or S-values for all potential sites of cleavage. The predicted site of signal peptidase cleavage occurs where the S-value is the highest (von Heijne, 1986). A computer program for calculating S-values (SIGSEQ2) has previously been described in another report (Folz and Gordon, 1987). It is interesting to note that an older method for predicting the site of signal peptidase cleavage which does not use a weight matrix (von Heijne, 1983) predicts cleavage to occur after Gly¹⁸↓ for all mutants listed above, with the exception of Ala²⁰, in which case cleavage would be predicted to occur after Ala²⁰↓.

^d The percent processing was calculated after laser densitometric tracing of autoradiographs. This method of determining the extent of processing of a mutant protein does not allow us to discriminate between the cleavage after residues 15, 18, or 20; rather, it integrates proteolytic cleavages at all observed sites. Mutants are categorized into three groups based on their processing efficiencies: those with the highest processing efficiency (+++), those with the lowest processing efficiency (+), and those that fall in between (++). To control for differences between experiments, the processing efficiencies for all mutants were normalized to the processing efficiency of preproapoA-II, which was determined for each set of reactions (e.g. processing of the Ala²⁰ mutant was 51% of that observed for preproapoA-II after 30 min. The absolute amount of processing of the wild-type preproprotein at 30 min is 35% (see Fig. 4). The data are accurate to $\pm 20\%$ (e.g. Ala²⁰ = $51 \pm 10\%$).

* The data for the Pro²⁰ mutant was confirmed in two independent experiments.

-2 of the Ring-necked pheasant prelysozyme c signal peptide which substitutes a Pro for a Leu⁻² shifts the site of signal peptide cleavage from Gly⁻¹ (Leu-Gly⁻¹) to the adjacent Pro (Pro-Gly). A bacterial hybrid preprotein (see Introduction) was recently shown to be cleaved by eukaryotic signal peptidase at two positions (Ser-Ser⁻¹-Asn-Ala⁻¹), with cleavage after Ser⁻¹ preferred (Garcia *et al.*, 1987).

Based on the limited number of examples of alternative/multiple co-translational cleavages, one can argue that cleavage by signal peptidase must be highly specific. What then are the features of a preprotein which define the site of signal peptide cleavage; or, to put it another way, what is the substrate specificity of eukaryotic signal peptidase? We demonstrate here that amino acids which occupy position -1 in the signal peptide of pre(Δ pro)apoA-II are critical in defining a good context for signal peptidase cleavage. For example, replacement of Ala²⁰ with Cys²⁰ results in >95% cleavage after Cys²⁰ and less than 5% cleavage after Gly¹⁸. Replacement of Ala²⁰ with Gly, Thr, Ser, or Pro results in cleavage after both Xaa²⁰ and Gly¹⁸. Replacement of Ala²⁰ with any of the charged (Asp²⁰, His²⁰, Arg²⁰), large hydrophobic (Leu²⁰, Ile²⁰, Val²⁰), or aromatic (Tyr²⁰) amino acids precludes cleavage after Xaa²⁰ and results in cleavage after Gly¹⁸. Evidence is also presented which suggests that cleavage can occur at a third position, Ser¹⁵, in certain sequence contexts (Leu²⁰, Ile²⁰), albeit at low efficiency (<5%). What do these results tell us about the role of position -1 in defining a good site for

signal peptidase cleavage? Fig. 7 attempts to correlate cleavage after Xaa²⁰ with two physical-chemical properties of the amino acid side chains for each pre(Δ pro)apoA-II mutant: their polarity and accessible surface area. From this analysis, we concluded that for any given polarity (e.g. Cys, Thr, Pro), the larger the accessible surface area of the side chain, the worse the substrate for signal peptidase cleavage. Also, for any given accessible surface area (e.g. Ala, Ser), the more polar the amino acid, the worse the substrate. Other physical-chemical characteristics including α -helix, β -strand, and turn propensities, various hydrophobicity scales, pI, and bulkiness were analyzed but no general conclusions could be made regarding the relationship between these indexes and cleavage site selection.

Our studies show the following preference of eukaryotic signal peptidase for cleavage after residue 20 (Xaa²⁰) of pre(Δ pro)apoA-II: Ala, Cys > Gly > Ser, Thr > Pro > Asn, Val, Ile, Leu, Tyr, His, Arg, Asp. No cleavage occurred after amino acids in the last group. From this study, it is not clear whether cleavage following these amino acids is prohibited or simply much less efficient compared to the Gly¹⁸ site. One possible method of differentiating among amino acids in this group would be to change the sequence of pre(Δ pro)apoA-II by substitution mutagenesis at positions other than -1 so as to suppress cleavage after Gly¹⁸.

The results obtained with the Gly²⁰ mutant further illustrate a dilemma concerning the interpretation of these data.

THIS PAGE BLANK (USPTO)

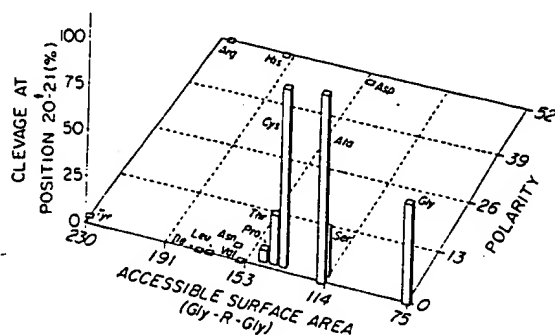


FIG. 7. Relationship between the physical-chemical properties of the 14 amino acids at residue 20 of pre(Δ pro)apoA-II and cleavage at this site. The height of each column represents the percent cleavage occurring after Xaa²⁰¹. The accessible surface area (Lee and Richards, 1971) of an amino acid side chain in the tripeptide Gly-R-Gly was determined by calculating the sum of the area on the surface of each atom present in the side chain of residue R, on each point of which the center of a solvent molecule (H_2O) can be placed in contact with this atom without penetrating any other atoms of the molecule (Chothia, 1976). The polarity index (which includes components of the dipole moment of the side chain and as well as any part due to the charge on an ionized group) attempts to relate the electric force due to the side chain acting on its immediate surroundings (Zimmerman *et al.*, 1968).

For example, does approximately equal cleavage after Gly¹⁸¹ and Gly²⁰¹ indicate that both sites are equally desirable (independent of other factors) or that for this construct the distance between the COOH-terminal residue of the hydrophobic core and the site of cleavage does not influence processing? It has been shown that a site-directed substitution mutation in bacterial prepenicillinase results in two distinct sites of cleavage. However, processing at these two sites occurs at different rates (Hayashi *et al.*, 1984). In this report, the processed proteins used for NH₂-terminal protein sequence analysis were all isolated after 90-min incubations. It is possible that a different ratio of cleavage after residues 18 and 20 would be observed with shorter incubations; although in some cases, generation of these data could be difficult given the overall efficiency of processing at early time points.

How does our grouping of amino acids at position -1 of the model preprotein reflect that found in a comparable position of naturally occurring preproteins? von Heijne (1986) analyzed 161 known eukaryotic preprotein sequences and tabulated the frequency of occurrence for all 20 amino acids for every position from -12 to +2. If one assumes that evolution optimizes biological processes (e.g. signal peptide cleavage), then one natural measure of eukaryotic substrate specificity is the frequency at which particular amino acids are represented at position -1. By assigning statistical weights for a given amino acid at a given position that are based on such frequency analysis, von Heijne (1986) was able to generate a quantitative measure for predicting the most likely site of signal peptidase cleavage for any preprotein sequence. Our ranking of the best substrates at position -1 of pre(Δ pro)apoA-II is almost in agreement with the statistical weights von Heijne (1986) assigns each amino acid with two significant exceptions. First, Cys is given a weight about equal to that of Ser, whereas the data presented here suggest that Cys should be given a similar weight to Ala. Second, Thr is assigned a weight about one-third that of Ser, whereas the data presented above suggest Thr should be assigned the same weight as Ser. However, there may be reasons why Cys and Thr are under-represented in the preprotein data base at position -1, independent of the substrate specificity of signal peptidase. For example, Cys can form disulfide bonds which

stabilize protein structure, and this may be deleterious to protein translocation (Maher and Singer, 1986). Alternatively, we demonstrated that the kinetics of processing of the Cys²⁰ pre(Δ pro)apoA-II mutant was much slower than that of the Ala²⁰ pre(Δ pro)apoA-II mutant, suggesting that signal peptidase may have the same specificity for Cys as Ala, but the rate of cleavage is reduced. Selection pressure would thus tend to reduce the frequency of occurrence of Cys at position -1 independent of its substrate specificity. Obviously, if an amino acid such as Cys is under-represented in von Heijne's data base, application of his rules for predicting signal peptide cleavage sites would in turn underestimate the potential for cleavage after this residue in model mutant preproteins. One way of further evaluating the relative specificity of signal peptidase for cleavage after these 2 residues would be to replace Ala in other preproteins with Cys to determine if they are functionally interchangeable.

In summary, the observation that deleting the propeptide in preapoA-II redirects signal peptide cleavage has provided an interesting and novel model system for systematically studying structural features important in defining the site of signal peptide cleavage. We hope that several of the mutants described in this report (e.g. those with bidirected cleavage) will allow us to probe other features that may further contribute to a favorable context for signal peptidase cleavage.

Acknowledgments—We would like to thank Harold Sims for his help in preparing canine pancreatic microsomes and Mark Freeman for computer graphics assistance. The excellent technical support provided by the Protein Chemistry Facility (Department of Biological Chemistry, Washington University School of Medicine, St. Louis, MO) is gratefully acknowledged. Glen Hortin and Arnie Strauss provided stimulating discussion as well as critical reading of this manuscript.

REFERENCES

- Austen, B. M. (1979) *FEBS Lett.* 103, 308-312.
- Baker, R. K., Bentivoglio, G. P., and Lively, M. O. (1986) *J. Cell. Biochem.* 32, 193-200.
- Birnboim, H. C., and Doly, J. (1979) *Nucleic Acids Res.* 7, 1513-1523.
- Blobel, G., and Dobberstein, B. (1975) *J. Cell. Biol.* 67, 835-851.
- Burke, D. T., and Olson, M. V. (1986) *DNA (N. Y.)* 5, 325-332.
- Caulfield, M. P., Duong, L. T., and Rosenblatt, M. (1986) *J. Biol. Chem.* 261, 10953-10956.
- Chothia, C. (1976) *J. Mol. Biol.* 105, 1-14.
- Dev, I. K., Harvey, R. J., and Ray, P. H. (1985) *J. Biol. Chem.* 260, 5891-5894.
- Dierstein, R., and Wickner, W. (1986) *EMBO J.* 5, 427-431.
- Evans, E. A., Gilmore, R., and Blobel, G. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 581-585.
- Folz, R. J., and Gordon, J. I. (1986) *J. Biol. Chem.* 261, 14752-14759.
- Folz, R. J., and Gordon, J. I. (1987) *Biochem. Biophys. Res. Commun.* 146, 870-877.
- Garcia, P. D., Ghayeb, J., Inouye, M., and Walter, P. (1987) *J. Biol. Chem.* 262, 9463-9468.
- Ghayeb, J., Lunn, C. A., Inouye, S., and Inouye, M. (1985) *J. Biol. Chem.* 260, 10961-10965.
- Gordon, J. I., Smith, D. P., Alpers, D. H., and Strauss, A. W. (1982) *J. Biol. Chem.* 257, 8418-8423.
- Gordon, J. I., Budelier, K. A., Sims, H. F., Edelstein, C., Scanu, A. M., and Strauss, A. W. (1983a) *J. Biol. Chem.* 258, 14054-14059.
- Gordon, J. I., Sims, H. F., Lentz, S. R., Edelstein, C., Scanu, A. M., and Strauss, A. W. (1983b) *J. Biol. Chem.* 258, 4037-4044.
- Hanahan, D. (1985) *DNA Cloning Technology*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Hayashi, S., Chang, S. Y., Chang, S., and Wu, H. C. (1984) *J. Biol. Chem.* 259, 10448-10454.
- Hortin, G., and Boime, I. (1981) *Cell* 24, 453-461.
- Innis, M. A., Tokunaga, M., Williams, M. E., Loranger, J. M., Chang, S. Y., Chang, S., and Wu, H. C. (1984) *Proc. Natl. Acad. Sci. U. S. A.* 81, 3708-3712.
- Inouye, S., Duffaud, G., and Inouye, M. (1986) *J. Biol. Chem.* 261, 10970-10975.

THIS PAGE BLANK (USPTO)

- Jackson, R. C., and White, W. R. (1981) *J. Biol. Chem.* 256, 2545-2550
- Kaiser, C. A., Preuss, D., Grisafi, P., and Botstein, D. (1987) *Science* 235, 312-317
- Kreibich, G., Czako-Graham, M., Grebenau, R. C., and Sabatini, D. D. (1980) *Ann. N. Y. Acad. Sci.* 343, 17-33
- Laemmli, U. K. (1970) *Nature* 227, 680-685
- Lee, B., and Richards, F. M. (1971) *J. Mol. Biol.* 55, 379-400
- Lehnhardt, S., Pollitt, S., and Inouye M. (1987) *J. Biol. Chem.* 262, 1716-1719
- Lingappa, V. R., Thiery-Devillers, A., and Blobel, G. (1977) *Proc. Natl. Acad. Sci. U. S. A.* 74, 2432-2436
- Lingappa, V. R., Chaidez, J., Spencer-Yost, C., and Hedgpeth, J. (1984) *Proc. Natl. Acad. Sci. U. S. A.* 81, 456-460
- Lively, M. O., and Walsh, K. A. (1983) *J. Biol. Chem.* 258, 9488-9495
- Maher, P. A., and Singer, S. J. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 9001-9005
- Mollay, C., Vilas, U., and Kreil, G. (1982) *Proc. Natl. Acad. Sci. U. S. A.* 79, 2260-2263
- Muller, M., Ibrahimi, I., Chang, C. N., Walter, P., and Blobel, G. (1982) *J. Biol. Chem.* 257, 11860-11863
- Perlman, D., and Halvorson, H. O. (1983) *J. Mol. Biol.* 167, 391-409
- Pluckthun, A., and Knowles, J. R. (1987) *J. Biol. Chem.* 262, 3951-3957
- Pollitt, S., Inouye, S., and Inouye, M. (1986) *J. Biol. Chem.* 261, 1835-1837
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* 74, 5463-5467
- Schauer, I., Emr, S., Gross, C., and Schekman, R. (1985) *J. Cell Biol.* 100, 1664-1675
- Stader, J., Benson, S. A., and Silhavy, T. J. (1986) *J. Biol. Chem.* 261, 15075-15080
- Stern, J. B., and Jackson, R. C. (1985) *Arch. Biochem. Biophys.* 237, 244-252
- Strauss, A. W., Zimmerman, M., Boime, I., Ashe, B., Mumford, R. A., and Alberts, A. W. (1979) *Proc. Natl. Acad. Sci. U. S. A.* 76, 4225-4229
- Thomas, K. A., Silverman, R. E., Jeng, I., Baglan, N. C., and Bradshaw, R. A. (1981) *J. Biol. Chem.* 256, 9147-9155
- von Heijne, G. (1983) *Eur. J. Biochem.* 133, 17-21
- von Heijne, G. (1984) *J. Mol. Biol.* 173, 243-251
- von Heijne, G. (1985) *J. Mol. Biol.* 184, 99-105
- von Heijne, G. (1986) *Nucleic Acids Res.* 14, 4683-4690
- Walter, P., and Blobel, G. (1983) *Methods Enzymol.* 96, 84-93
- Watson, M. E. E. (1984) *Nucleic Acids Res.* 12, 5145-5164
- Watts, C., Wickner, W., and Zimmermann, R. (1983) *Proc. Natl. Acad. Sci. U. S. A.* 80, 2809-2813
- Weisman, L. S., Krummel, B. M., and Wilson, A. C. (1986) *J. Biol. Chem.* 261, 2309-2313
- Wolfe, P. B., Wickner, W., and Goodman, J. M. (1983) *J. Biol. Chem.* 258, 12073-12080
- Wu, H. C. (1986) in *Protein Compartmentalization* (Strauss, A. W., Boime, I., and Kreil, G., eds) pp. 33-60, Springer-Verlag, New York
- Zimmerman, J. M., Eliezer, N., and Simha, R. (1968) *J. Theor. Biol.* 21, 170-201

THIS PAGE BLANK (USPTO)

Alterations in the cleavage site of the signal sequence for the secretion of human lysozyme by *Saccharomyces cerevisiae*

Hitoshi Nagahora, Hirofumi Fujisawa and Yoshifumi Jigami

Bioorganic Chemistry Division, National Chemical Laboratory for Industry, Higashi 1-1, Tsukuba, Ibaraki 305, Japan

Received 27 July 1988

The amino acids corresponding to the cleavage site of a hybrid preprotein containing a chicken lysozyme signal and a mature portion of human lysozyme were altered. The processing of mutant signals of -3Pro and -3Asp/-1Ala decreased remarkably, while that of -2Pro was 75% of that of the native signal. The major cleavage site of -3Pro was the same as that of the native signal, but that of the -2Pro and -3Asp/-1Ala signals was shifted one residue closer to the N-terminal side than the original site. The cleavage of the -2Pro signal, which was identical to the native processing of pheasant prelysozyme, suggested that the signal peptidases in yeast and bird are similar.

Signal peptidase; Yeast; Cleavage site; Lysozyme signal; (Chicken)

1. INTRODUCTION

The signal peptide is composed of three structurally and functionally distinct regions: a basic N-terminal region, a central hydrophobic region, and a more polar C-terminal region defining the cleavage site. To confirm the cleavage specificity in yeast signal peptidase, we have engineered the cleavage site of a hybrid preprotein consisting of a CLY signal and a mature HLY, because the CLY signal contains common features described above and is functional (forms a correctly processed mature HLY) in yeast [1].

Von Heijne proposed the (-3, -1) rules according to the statistical studies on signal sequences [2,3]; the residue at position -1 must be small and the residue at position -3 must not be aromatic, charged, or large and polar, and proline

must be absent from positions -3 to +1 with some exceptions. In this paper we describe the correlation between the processing efficiency and the (-3, -1) rules, and the shifts in the cleavage site of some mutant signals in yeast.

2. MATERIALS AND METHODS

2.1. Materials

Restriction enzymes and T₄ ligase were purchased from Toyobo or Takara Shuzo and used under the conditions as recommended by the suppliers. *Micrococcus lysodeikticus* cells and authentic HLY were obtained from Sigma. *Saccharomyces cerevisiae* KK4 (α leu2 ura3 his1 or 3 trp1 gal80) [4] was used as the host strain. Synthetic media [5], supplemented with an amino acid mixture (20-375 μ g/ml) lacking leucine and with adenine sulfate (20 μ g/ml) and uracil (20 μ g/ml) [1] in addition to 2% glucose, were used to cultivate *S. cerevisiae*.

2.2. DNAs

The cassette mutagenesis methods [6] were used to construct a series of pFJ300 plasmids by replacing a part of the original DNA segments with double-stranded DNA oligomers which contain mutations at the residues in positions -1 and/or -3. The site-directed mutagenesis [7] was also performed to construct a series of pNJ300 plasmids by using the Amersham system. The multi-copy plasmid used to express HLY in yeast was the same as described in [1] except that the *GAL10* promoter was replaced with the cloned *ENO1* promoter [8].

Correspondence address: Y. Jigami, National Chemical Laboratory for Industry, Higashi 1-1, Tsukuba, Ibaraki 305, Japan

Abbreviations: HLY, human lysozyme; CLY, chicken lysozyme; aa, amino acid; preHLY, precursor HLY

THIS PAGE BLANK (USPTO)

2.3. Assay of processing efficiency

The efficiency of signal peptidase to cleave preHLY was evaluated either by the initial decrease in preHLY labeled with [³⁵S]methionine or by the ratio of mature HLY/(mature HLY + preHLY) measured by Western blotting. The lysozyme activity in the medium was assayed by monitoring the lysis of *M. lysodeikticus* cells at 450 nm.

2.4. N-terminal sequence

Secreted HLYs obtained from 2.5 l cultures were purified as described in [1] and were applied to an automatic protein sequencer (model 477A, Applied Biosystems). The aa sequences of the 5 N-terminal residues were quantitatively analysed by measuring the amounts of phenylthiohydantoin amino acids to determine the cleavage sites of mutant signals.

3. RESULTS

3.1. Secretion of HLY

The native CLY signal contains 18 aas like a yeast melibiase signal [9]. We replaced (i) Gly at position -1 with a larger aa (Ala, Ser, Thr or Val), (ii) Ala at position -3 with a rare locating aa (Val, Asp or Pro), (iii) Leu at position -2 with the β -turn forming aa (Pro). The effects of signal mutation on HLY secretion are summarized in table 1. The amounts of secreted HLY were almost the

same in mutant signals containing -3Val, and were about half in mutant signals containing -1Ala, -1Ser, -3Val/-1Ala, and -2Pro, as those in the native signal. No lysozyme activity was detected in the other mutant signals without any concentrations of media under the assay conditions.

3.2. Processing of mutant preHLY

To evaluate the efficiency of preHLY processing, the conversion of preHLY containing mutations at position -1 to mature HLY was examined by immunoprecipitation. The amount of [³⁵S]methionine labeled preHLY decreased during the chase period in native and mutant signals, while that of mature HLY also decreased after a slight increase (fig.1A). As the processing was thought to be reflected better in the decrease of preHLY than in the increase of mature HLY, the decay curves of the mutant preHLYs were compared. It was found that the initial decreasing rate of preHLY decreased with the increase in the bulkiness of the aa at position -1 (fig.1B).

The correlation of the decay of preHLY

Table 1
Effect of mutant signal peptides on HLY secretion

Class ^a	Plasmid ^a	HLY activity ^b (units/ml)	Relative ratio	Processing ^c (M/M + P, %)	Relative ratio
WT	pFJ1053	64	1.00	68	1.00
I	pNJ301 (-1A)	31	0.48	68	0.85
	pNJ302 (-1S)	26	0.41	63	0.79
	pFJ342 (-1T)	0	0	41	0.51
	pFJ330 (-1V)	0	0	31	0.39
II	pFJ335 (-3V)	62	0.97	78	0.98
	pFJ371 (-3V/-1A)	31	0.48	78	0.98
	pFJ327 (-3V/-1V)	0	0	17	0.21
III	pFJ338 (-3V/-1D)	0	0	26	0.33
	pNJ303 (-1D)	0	0	6	0.08
	pFJ340 (-3D/-1A)	0	0	36	0.45
IV	pNJ305 (-2P)	26	0.41	60	0.75
	pNJ304 (-3P)	0	0	27	0.34

^a WT, wild type preHLY; I, -1 mutants containing a different size of aa; II, -3 and/or -1 mutants containing a bulky aa; III, -3 and/or -1 mutants containing a charged aa; IV, -2 or -3 mutant containing a β -turn forming aa. aa are given as capital letters

^b The lysis of 0.9 ml of *M. lysodeikticus* cells (0.15 mg/ml) was measured by adding 0.1 ml of culture supernatant after 4 days of cultivation. One unit of HLY is defined as the amount of enzyme which decreases 0.001 of A_{450} per min at 25°C. In this assay condition, the specific activity of an authentic HLY (purified from human milk, Sigma) was 130000 units/mg

^c Determined by densitometric scanning of the autoradiogram from the Western blot gel and expressed as a percent of mature HLY/(mature HLY + preHLY)

THIS PAGE BLANK (USPTO)

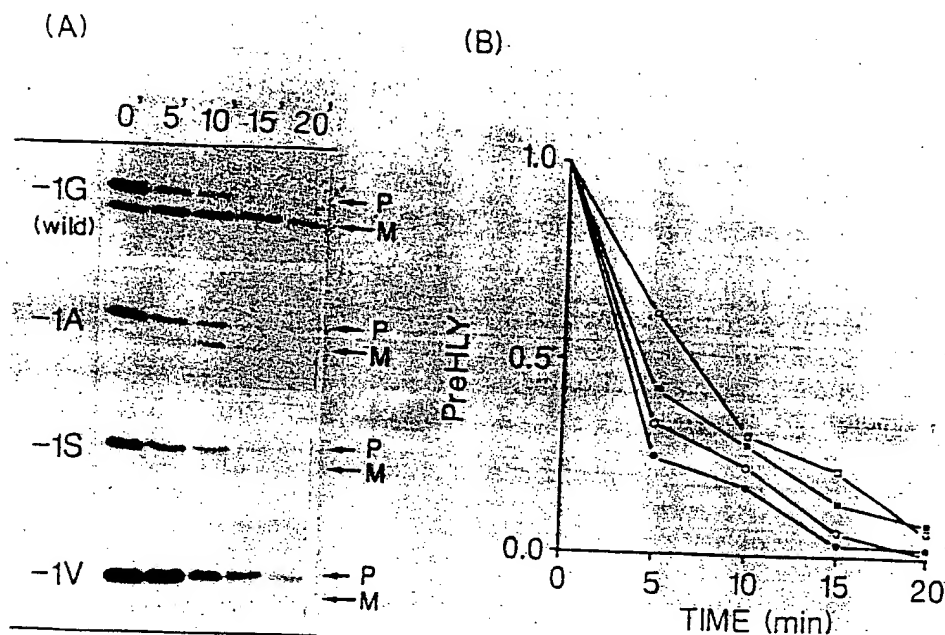


Fig.1. (A) Immunoprecipitation and SDS-polyacrylamide gel electrophoresis of mutant preHLYs. The KK4 cells transformed with the plasmid containing either wild type or mutant preHLY genes were spheroplasted by Zymolyase and labeled with [35 S]methionine (60 μ Ci) for 10 min. Immediately after initiation of the chase period unlabeled methionine (40 mM) and 50 mM cycloheximide were added. Aliquots were taken at various times and immunoprecipitated with HLY antiserum. Immunoprecipitates were analysed by autoradiography after gel electrophoresis. P, preHLY; M, mature HLY; -1G, -1A, -1S and -1V showed the signals containing Gly (wild type), Ala, Ser and Val at position -1, respectively. (B) Decrease of preHLYs during the pulse chase period. The intensity of the preHLY band (P) in A was measured by densitometry and the relative values at each time to that at zero time are shown. (●) WT; (○) -1Ala; (■) -1Ser; (□) -1Val.

measured by the pulse chase method to the amount of mature HLY was also confirmed by measuring the amounts of mature HLY (M) and preHLY (P) in the exponentially grown cells from Western blots (not shown). The processing efficiency (the ratio of M/M+P) is summarized in table 1. The efficiency is remarkably low (6–41%) in the mutant signals which contain a statistically prohibited or a rare locating aa at position -1 or -3, and it is relatively high (75%) in the -2Pro mutant.

Fig.2. The signal cleavage sites in various preproteins. Amino acids are shown as capital letters and the positions of replaced amino acids are shown by asterisks. The cleavage sites and the ratios of each cleavage are shown by arrows and percentages, respectively. WT, wild type preHLY containing a native CLY signal; -3P, -3Pro mutant; -2P, -2Pro mutant; -3D/-1A, -3Asp/-1Ala mutant; PLY, pheasant prelysozyme; SUC2, yeast preinvertase; SUC2 (-1V), mutant preinvertase containing Val at position -1.

	-18	M	R	S	L	L	I	
	-12	L	V	L	C	F	L	100%
WT	-6	P	L	A	A	L	G	K V
								100%
-3P		P	L	A	*	P	L	G K V
								88% 12%
-2P		P	L	A	A	*	P	G K V
								70% 30%
-3D/-1A		P	L	A	D	L	A	K V
PLY		P	L	A	A	P	G	K V
SUC2		A	A	K	I	S	A	S M
SUC2 (-1V)		A	A	K	I	S	V	S M

THIS PAGE BLANK (USPTO)

3.3. Cleavage site determination

HLYs processed from the preHLYs containing such statistically prohibited signals as -3Pro, -2Pro, and -3Asp/-1Ala, were purified to confirm their cleavage sites. The unique site identical to that of the native signal was found in the -3Pro signal, suggesting the acceptance of Pro at position -3 in yeast. Heterologous cleavages were found in -2Pro and -3Asp/-1Ala signals (fig.2). The major cleavage site of the -2Pro signal (88%) and the -3Asp/-2Ala signal (70%) was shifted one aa closer to the N-terminal side than that of the native signal, although the minor site (12-30%) was identical to that of the native signal.

4. DISCUSSION

The same shifts in the main cleavage site were found in -2Pro and -3Asp/-1Ala signals, despite the great differences in amounts of HLY secreted (265 µg/l for -2Pro and 1.7 µg/l for -3Asp/-1Ala mutants). The -3Pro signal commanded a poor secretion of HLY (8.8 µg/l), but was processed at the same site as the native signal. The computer program, SIGSEQ2 [11], based on a statistic weight-matrix method [12], was applied for the mutant signals to evaluate the most probable site of cleavage. The actual major site of cleavage for -2Pro and -3Asp/-1Ala signals has the highest processing probability, but that for the -3Pro signal had the third highest score (not shown), suggesting that this method is insufficient to predict the processing site of the statistically rare aa sequences.

The cleavage site of the -2Pro mutant was identical to that of pheasant prelysozyme, which contains the same signal sequence as the -2Pro signal [13]. This suggests that the signal peptidase is similar in yeast and bird and prefers

the -3Ala/-1Pro sequence to the native -3Ala/-1Gly sequence. The -3Asp/-1Ala mutant containing a negatively charged aa at position -3 resembles (with respect to the poor processing and the shift in the cleavage site [14], although the sites of cleavage are different (fig.2)) the yeast mutant preinvertase containing a bulky aa (Val) at position -1 instead of the native Ala. It is probable that the signal peptidase is able to choose a suitable cleavage site when the original sequence is not fit to cleave.

Acknowledgement: We are grateful to Dr Rodney J. Folz for providing the computer program, SIGSEQ2, to evaluate the cleavage site.

REFERENCES

- [1] Jigami, Y., Muraki, M., Harada, N. and Tanaka, H. (1986) *Gene* 43, 273-279.
- [2] Von Heijne, G. (1983) *Eur. J. Biochem.* 133, 17-21.
- [3] Von Heijne, G. (1986) *Nucleic Acids Res.* 12, 505-519.
- [4] Nogi, Y., Shimada, H., Matsuzaki, Y., Hashimoto, H. and Fukasawa, T. (1984) *Mol. Gen. Genet.* 195, 29-34.
- [5] Sherman, F., Fink, G.R. and Hicks, J.D. (1986) in: *Methods in Yeast Genetics*, p.164, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [6] Muraki, M., Jigami, Y., Morikawa, M. and Tanaka, H. (1987) *Biochim. Biophys. Acta* 911, 376-380.
- [7] Taylor, J.W., Ott, J. and Eckstein, F. (1985) *Nucleic Acids Res.* 13, 8764-8785.
- [8] Uemura, H., Shiba, T., Paterson, M., Jigami, Y. and Tanaka, H. (1986) *Gene* 45, 67-75.
- [9] Summer Smith, M., Bozzato, R.P., Skipper, N., Davies, R.W. and Hopper, J.E. (1985) *Gene* 36, 333-340.
- [10] Laemmli, U.K. (1970) *Nature* 227, 680-685.
- [11] Folz, R.J. and Gordon, J.I. (1987) *Biochem. Biophys. Res. Commun.* 147, 870-877.
- [12] Von Heijne, G. (1986) *Nucleic Acids Res.* 14, 4683-4690.
- [13] Weisman, L.S., Krummel, B.M. and Wilson, A.C. (1986) *J. Biol. Chem.* 261, 2309-2313.
- [14] Schauer, I., Emr, S., Gross, C. and Schekman, R. (1985) *J. Cell Biol.* 100, 1664-1675.

THIS PAGE BLANK (USPTO)

Defective *Escherichia coli* signal peptides function in yeast

O. Pines,¹ C. A. Lunn² and M. Inouye^{3*}

Department of Biochemistry, State University of New York at Stony Brook, Stony Brook, New York 11794, USA.

Summary

To investigate structural characteristics important for eukaryotic signal peptide function *in vivo*, a hybrid gene with interchangeable signal peptides was cloned into yeast. The hybrid gene encoded nine residues from the amino terminus of the major *Escherichia coli* lipoprotein, attached to the amino terminus of the entire mature *E. coli* β -lactamase sequence. To this sequence were attached sequences encoding the non-mutant *E. coli* lipoprotein signal peptide, or lipoprotein signal peptide mutants lacking an amino-terminal cationic charge, with shortened hydrophobic core, with altered potential helicity, or with an altered signal-peptide cleavage site. These signal-peptide mutants exhibited altered processing and secretion in *E. coli*. Using the GAL10 promoter, production of all hybrid proteins was induced to constitute 4–5% of the total yeast protein. Hybrid proteins with mutant signal peptides that show altered processing and secretion in *E. coli*, were processed and translocated to a similar degree as the non-mutant hybrid protein in yeast (approximately 36% of the total hybrid protein). Both non-mutant and mutant signal peptides appeared to be removed at the same unique site between cysteine 21 and serine 22, one residue from the *E. coli* signal peptidase II processing site. The mature lipo- β -lactamase was translocated across the cytoplasmic membrane into the yeast periplasm. Thus the protein secretion apparatus in yeast recognizes the lipoprotein signal sequence *in vivo* but displays a specificity towards altered signal sequences which differs from that of *E. coli*.

Introduction

Protein translocation systems in eukaryotic and

prokaryotic cells share a number of common features. Both require the signal peptide, an amino acid extension attached to the amino terminus of the protein destined for translocation. Analysis of the sequences of known signal peptides (Watson, 1984) show that the peptides from both systems, while different in amino acid sequence, share similar structural characteristics. The amino terminus of the signal peptide frequently contains several positively charged amino acids. This is followed by a 14–20 amino acid hydrophobic domain. Together, these residues are proposed to form a hydrophobic loop in the membrane (Inouye and Halegoua, 1980).

Amino acids immediately adjacent to the site of cleavage of the signal peptide constitute a third domain of this sequence. The relative importance of these structural domains has been tested in bacteria using a large number of defined signal-peptide mutants. The major *E. coli* lipoprotein signal sequence, which is required for secretion of lipoprotein, has been thoroughly analysed in this way (Pollitt and Inouye, 1986). In eukaryotes, the same approach has been limited by the small number of such defined mutant signal peptides isolated.

In order to allow a comparative analysis of eukaryotic and prokaryotic secretion systems, we have expressed a prokaryotic protein with exchangeable signal peptides in yeast. Previous reports from this laboratory have dealt with the expression in *E. coli* of a protein containing the major lipoprotein signal sequence plus nine amino acids attached to the mature portion of β -lactamase. This hybrid protein, when expressed in *E. coli*, is post-translationally modified, processed, and localized to the bacterial outer membrane (Ghrayeb and Inouye, 1984). Exchange of the non-mutant lipoprotein signal peptide with signal-peptide mutants reveals altered processing and secretion (Lunn and Inouye, 1987). In this study, we have expressed both non-mutant and mutant hybrid lipo- β -lactamases in yeast. This prokaryotic protein was found to be processed and translocated into the yeast periplasm regardless of the lesion within the lipoprotein signal peptide.

Results

Cloning of lipo- β -lactamase under a GAL10 promoter

A GAL10 expression system was chosen for the production of the hybrid protein lipo- β -lactamase in yeast. The shuttle

Received 16 July, 1987; revised 14 November, 1987. Present addresses:

¹Department of Molecular Biology, Hadassah Medical School, The Hebrew University, Jerusalem, Israel; ²Department of Molecular Biology, Schering Corporation, 60 Orange Street, Bloomfield, NJ 07003, USA; ³Department of Biochemistry, Robert Wood Johnson Medical School, at Rutgers, UMDNJ, 675 Hoes Lane, Piscataway, NJ 08854–5635, USA.

*For correspondence.

THIS PAGE BLANK (USPTO)

vector YEp51 (Broach *et al.*, 1983) used in this study contains, in addition to the strong inducible GAL10 promoter, a termination of transcription from the yeast 2 μ m circle. It also contains origins of replication and selective marker(s) required for maintaining the vector as an extra-chromosomal plasmid in both yeast and *E. coli*. The source of the hybrid gene was plasmid pJG311, which was used previously for expression of lipo- β -lactamase in *E. coli* (Ghrayeb and Inouye, 1984). The strategy for cloning of the lipo- β -lactamase gene into the expression vector is shown in Fig. 1. A 1.5-kb *Xba*I/*Bam*HI fragment from plasmid pJG311 was inserted between the promoter and the termination site of YEp51 using the *Sal*I and *Bam*HI sites. The lipo- β -lactamase gene in YEp51 contains a 16-base-pair sequence upstream of the initiating ATG from the prokaryotic gene in which there are no other initiation signals. The YEp51 plasmid containing the lipo- β -lactamase gene (designated pCOP1) was then used to transform a leucine-requiring yeast strain. Selected transformants analysed by Southern blot hybridization contained the lipo- β -lactamase gene (data not shown).

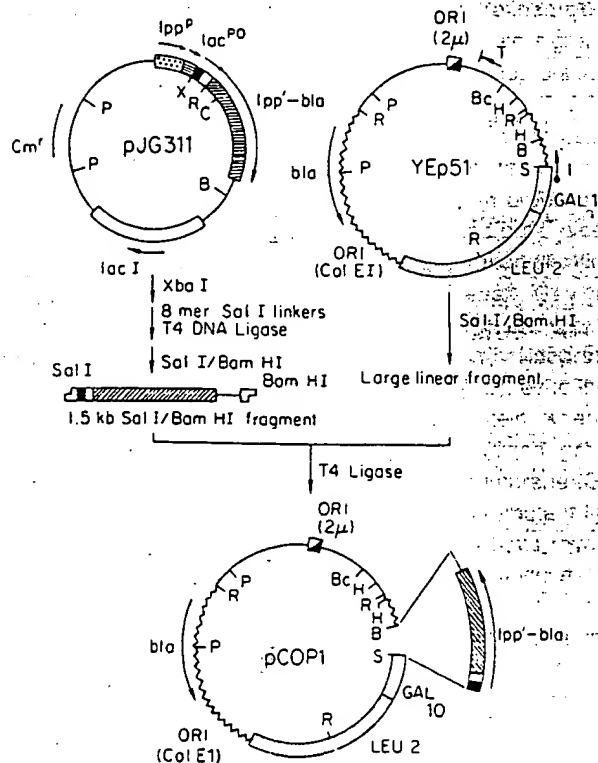


Fig. 1. Schematic diagram of pCOP1 construction (not to scale). The coding sequence of *E. coli* lipo- β -lactamase from plasmid pJG311 was inserted between the GAL10 promoter and the transcription termination site of YEp51. Abbreviations: bla, β -lactamase; lpp'bla, lipo- β -lactamase; Cm^r, chloramphenicol-resistance marker; B, *Bam*HI; Bc, *Bcl*I; C, *Hinc*II; H, *Hind*III; P, *Pst*I; R, *Eco*RI; S, *Sal*I; X, *Xba*I.

Expression of lipo- β -lactamase in yeast

To examine the expression of lipo- β -lactamase, yeast cells carrying pCOP1 were grown in medium containing either galactose or dextrose as the carbon source. Exponentially growing cells were labelled with [³⁵S] methionine and cell-free extracts were prepared and analysed by SDS polyacrylamide gel electrophoresis. Aliquots of each extract were also immunoprecipitated with β -lactamase antiserum in order to identify β -lactamase cross-reacting material. Two protein species were specifically induced by galactose in pCOP1-containing cells (Fig. 2, lane 4). Both species were immunoprecipitated with β -lactamase antiserum (Fig. 2, lane 8) and were absent from cells grown on dextrose or from cells harbouring the control plasmid YEp51 (Fig. 2, lanes 1–3 and 5–7). The two inducible bands corresponded in size to the precursor and mature forms of lipo- β -lactamase in *E. coli*. Quantitation of the relative amounts of radioactivity present in the two bands (Fig. 2, lane 4) shows that these two species constitute roughly 4–5% of the total yeast soluble protein pool. Quantitation of the two immunoprecipitated bands (Fig. 2, lane 8) suggests that 30–40% of the hybrid gene product has been processed to a mature form of lipo- β -lactamase, with the remainder being in precursor form.

It is important to point out that vector YEp51 harbours a β -lactamase gene controlled by the original prokaryotic promoter. The production of β -lactamase from this prokaryotic promoter is negligible in yeast and should not be confused with the production of lipo- β -lactamase from the GAL10 promoter since: (a) in a strain containing the original YEp51 no β -lactamase cross-reacting material was detected by the polyacrylamide gel electrophoresis analysis described above (Fig. 2, lanes 1, 2, 5 and 6); (b) the β -lactamase-like bands were galactose-inducible, indicating that the yeast GAL10 promoter is responsible for directing production of these proteins; and (c) the apparent sizes of these products correspond to the sizes of lipo- β -lactamase gene products and not to the sizes of the β -lactamase gene products (which are 13 amino acids smaller).

Activity of lipo- β -lactamase in yeast

It was considered to be worth investigating whether lipo- β -lactamase synthesized in yeast possessed enzymatic activity. Two semi-quantitative assays were used to determine β -lactamase activity in culture lysates. Cell-free extracts from induced and uninduced cultures were spotted onto solid agar medium containing 10–20 μ g ml⁻¹ ampicillin. After incubation at room temperature for 5 min, an ampicillin-sensitive *E. coli* strain was overlaid onto the plates in soft agar, and the plates were incubated at 37°C overnight. Extracts from induced yeast cells harbouring pCOP1 potentiated growth of the indicator strain, while

THIS PAGE BLANK (USPTO)

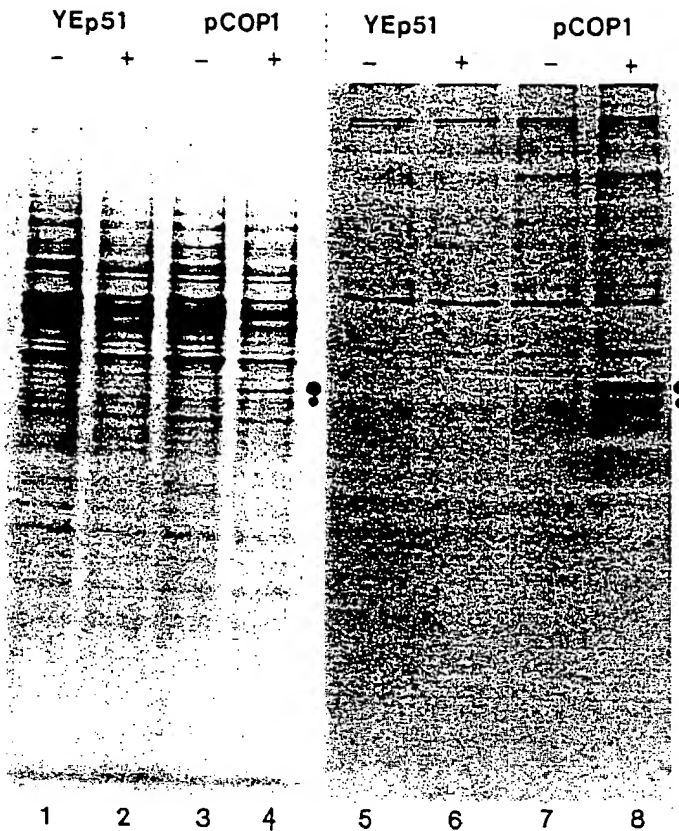


Fig. 2. Production of the lipo- β -lactamase in yeast. *S. cerevisiae* DMM1-15A harbouring plasmid pCOP1 or YE51 were grown to 1×10^7 cells ml^{-1} in dextrose medium (uninduced cultures) or galactose medium (induced cultures) and labelled for 1 h with [^{35}S]-methionine. Aliquots of extracts prepared from the labelled cells were immunoprecipitated with antiserum to β -lactamase and then analysed by SDS polyacrylamide gel electrophoresis. Total (lanes 1–4) and immunoprecipitated extracts (lanes 5–8) of induced (+) and uninduced (–) cultures harbouring either pCOP1 (lanes 3, 4, 7, and 8) or YE51 (lanes 1, 2, 5 and 6) are presented. Symbols: the large dot shows the position of prolipo- β -lactamase and the small dot shows the position of mature lipo- β -lactamase.

extracts from induced cells carrying the vector YE51 allowed no growth. The diameter of the growth ring produced by the induced pCOP1 cell extracts after 10–16 h was at least ten times larger than the diameter for uninduced cell extracts.

The method of Tai, Zyk and Citri (Tai *et al.*, 1985) was used to determine whether the β -lactamase activity was

associated with both forms of the hybrid protein. Extracts from induced cells harbouring pCOP1 and YE51 were subjected to SDS polyacrylamide gel electrophoresis. The SDS was removed from the gels and the β -lactamase was renatured by washing the gel in the presence of ampicillin. β -lactamase activity was then determined with starch-iodine paper, as described in the *Experimental procedures*

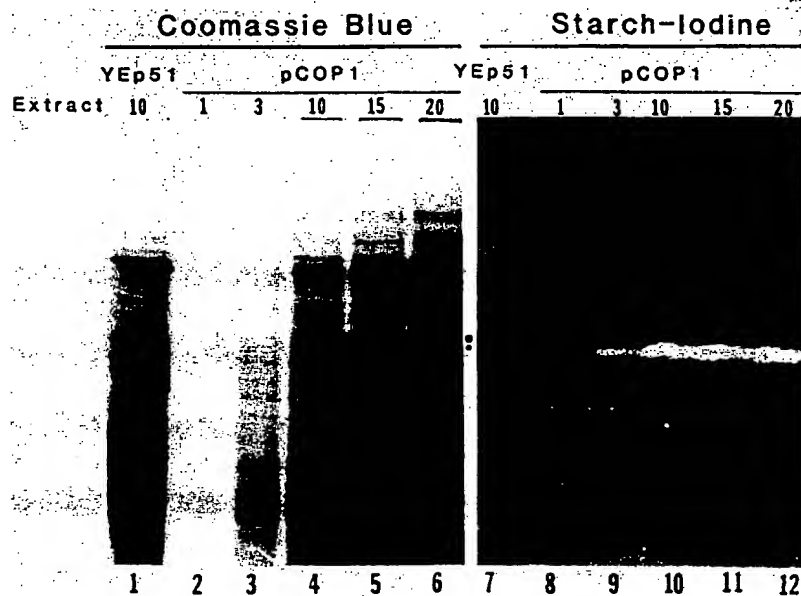


Fig. 3. Activity of lipo- β -lactamase in yeast. Total cell extracts were prepared from induced cultures harbouring pCOP1 or YE51, subjected to SDS polyacrylamide gel electrophoresis, and then assayed for β -lactamase activity using starch-iodine paper (as described in *Experimental procedures*).

Panel A: Coomassie-blue-stained gel after assay, with 10 μl of a YE51 culture extract and 1 μl , 3 μl , 10 μl , 15 μl and 20 μl of a pCOP1 culture extract. Panel B: Starch-iodine paper after the assay with the gel in Panel A. Symbols: the large dot shows the position of prolipo- β -lactamase and the small dot shows the position of mature lipo- β -lactamase.

THIS PAGE BLANK (USPTO)

section. Figure 3 shows that β -lactamase activity was detected using 1 μ l of an induced pCOP1 cell extract (lanes 2 and 8), whereas no activity was observed using 10 μ l of a similar YEp51 cell extract (lanes 1 and 7). All the observed activity migrated with the smaller of the two immunoprecipitated protein bands, which corresponds to the processed form of lipo- β -lactamase. These results are consistent with the observation that only the processed form of β -lactamase is active (Roggenkamp *et al.*, 1985) and indicates that processed lipo- β -lactamase with enzymatic activity can be produced in yeast.

Processing of lipo- β -lactamase

Previous studies showed that in *E. coli* the hybrid protein lipo- β -lactamase was cleaved at residue 20 when processed by the lipoprotein-specific signal peptidase II (Ghrayeb and Inouye, 1984), or at residue 25 when processed by the broadly specific signal peptidase I (Ghrayeb *et al.*, 1985). In order to determine the site at which lipo- β -lactamase is processed in yeast, cells harbouring pCOP1 were labelled with [3 H]-isoleucine. A soluble fraction prepared from these cells was then immunoprecipitated with antiserum to β -lactamase, and the immunoprecipitates were subjected to SDS polyacrylamide gel electrophoresis. The band corresponding to the mature lipo- β -lactamase was then excised from a stained gel, and the labelled protein eluted and subjected to sequential Edman degradation. Figure 4 shows that radioactive isoleucine

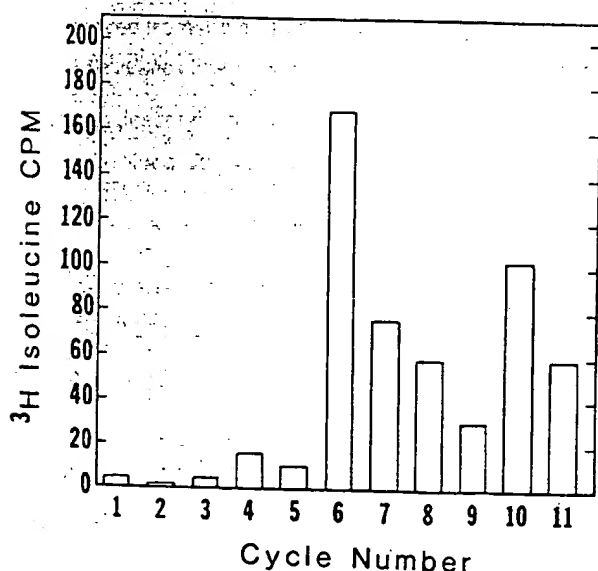


Fig. 4. Sequential Edman degradation of processed lipo- β -lactamase from yeast. Induced cultures harbouring pCOP1 were labelled for 3 h with 20 μ Ci ml^{-1} [3 H]-isoleucine. Total culture extracts were immunoprecipitated with β -lactamase antiserum and subjected to SDS polyacrylamide gel electrophoresis. The mature lipo- β -lactamase band was excised, eluted, and subjected to sequential Edman degradation (as described in Experimental procedures).

was detected in degradation cycles 6 and 10, indicating cleavage of the signal peptide between cysteine 21 and serine 22. Sequential Edman degradation following labelling of lipo- β -lactamase with serine revealed radioactive peaks in cycles 1 and 2 (data not shown), which is consistent with the result above.

Figure 2 shows that only part of the total prolipo- β -lactamase pool was processed *in vivo*. To confirm the precursor-product relationship between the two forms of prolipo- β -lactamase, induced cells were subjected to a pulse-chase experiment. Exponentially growing yeast cells harbouring pCOP1 were pulse-labelled for 1 min with 40 μ Ci ml^{-1} [35 S]-methionine and this was followed by the addition of 0.5 mg ml^{-1} non-radioactive methionine. At specified times, aliquots were removed, and cell metabolism stopped by dilution into formaldehyde. Figure 5 shows that the presumed precursor of lipo- β -lactamase was labelled exclusively after incubation for 1 min. The radioactive protein was then chased into the smaller lipo- β -lactamase species, reaching a maximum at 10 min after the addition of non-radioactive methionine. The finding that only 30–40% of the lipo- β -lactamase is secreted in yeast is similar to the partial secretion achieved with some heterologous eukaryotic proteins directed by yeast signal sequences (e.g. Smith *et al.*, 1985).

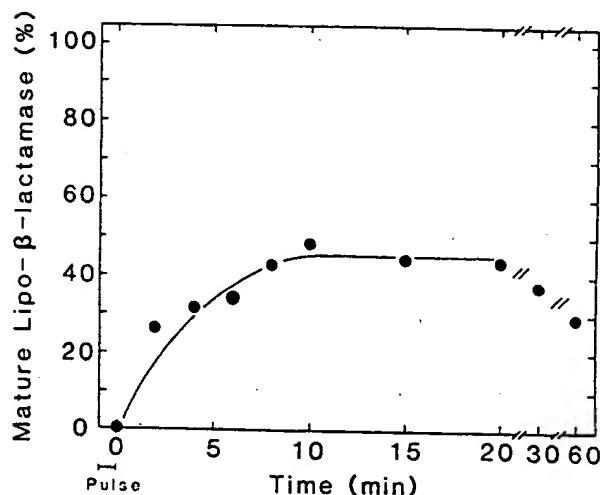


Fig. 5. Rate of prolipo- β -lactamase processing. Cells harbouring pCOP1 were grown in galactose medium minus methionine, pulse-labelled for 1 min with the addition of 40 μ Ci ml^{-1} [35 S]-methionine, and then chased with non-radioactive methionine (0.5 mg ml^{-1}). At specified times, aliquots were removed, cell extracts were prepared, immunoprecipitated with β -lactamase antiserum and analysed by SDS polyacrylamide gel electrophoresis. The relative amounts of prolipo- β -lactamase and lipo- β -lactamase were determined densitometrically from autoradiograms.

Localization of lipo- β -lactamase products in the cell

It was important to determine the subcellular localization of the lipo- β -lactamase gene products in yeast to implicate removal of the signal sequence with a secretion event.

THIS PAGE BLANK (USPTO)

Induced cells harbouring pCOP1 were first labelled for 3 h with [35 S]-methionine. Zymolyase-20T was then used to form spheroplasts which were collected and lysed, allowing fractionation of the cells into periplasmic, cytoplasmic and total membrane fractions. To monitor the fractionation procedure, we assayed glyceraldehyde-3-phosphate dehydrogenase, which is a cytoplasmic enzyme (Schekman, 1982), and α -galactosidase, which is a periplasmic enzyme (Tschopp *et al.*, 1984). Using the same conditions described in Fig. 2, over 90% of the glyceraldehyde-3-phosphate dehydrogenase activity was associated with the cytoplasmic fraction whereas about 90% of the α -galactosidase was associated with the periplasmic fraction. Similarly, immunoprecipitated invertase was detected essentially in the periplasmic fraction. As shown in Fig. 6, the processed form of lipo- β -lactamase was localized almost exclusively to the periplasmic fraction (lanes 1 and 2), whereas the unprocessed form was found in equal amounts in the cytoplasmic and membrane fractions (lanes 3, 4, 5, and 6). These results indicate that mature lipo- β -lactamase is translocated into the yeast periplasm. We could not detect β -lactamase activity or the β -lactamase immunochemically in concentrated samples of culture supernatants, which indicates that β -lactamase is not released from the yeast periplasm into the medium.

Processing and secretion of mutant lipo- β -lactamase in yeast

Previous studies have analysed the effect of mutations in the lipoprotein signal sequence on its ability to direct secretion in *E. coli*. To examine the effect of these mutations on secretion of lipo- β -lactamase in yeast, a number of mutant genes were cloned into YEp51 by a strategy analogous to that illustrated in Fig. 1. The mutants used in this study are presented in Fig. 7. These include mutations in the basic, hydrophobic, and cleavage domains of

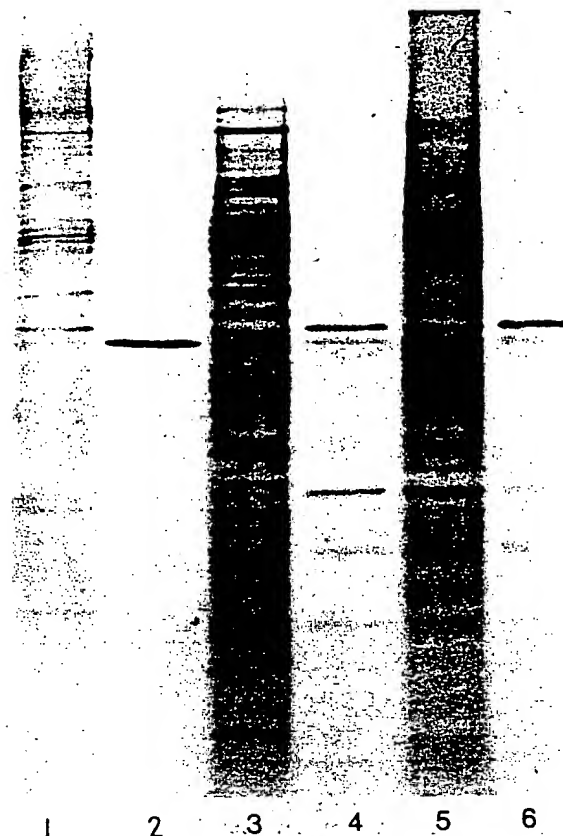


Fig. 6. Localization of lipo- β -lactamase. Cells harbouring pCOP1 were grown in galactose medium, and labelled for 3 h with [35 S]-methionine. The labelled cells were fractionated into periplasm (lanes 1 and 2), cytoplasm (lanes 3 and 4), and membrane (lanes 5 and 6) fractions (as described in *Experimental procedures*). Total (lanes 1, 3 and 5) and immunoprecipitated extracts (lanes 2, 4 and 6) were subjected to SDS polyacrylamide gel electrophoresis.

the lipoprotein signal peptide. All these mutants exhibited altered processing and secretion of lipo- β -lactamase in *E. coli* (Ghrayeb *et al.*, 1985; Lunn and Inouye, 1987). As

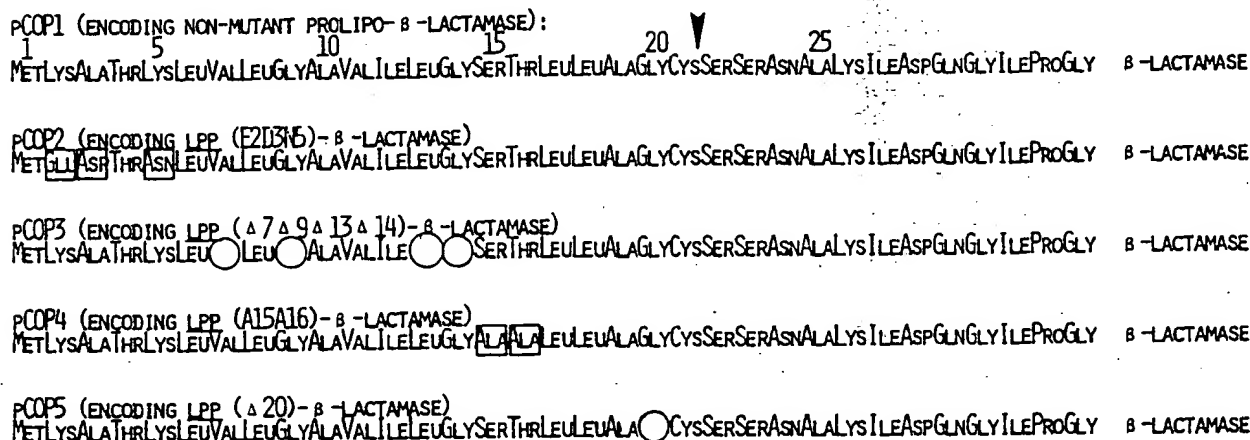


Fig. 7. Partial protein sequences of non-mutant and mutant hybrid proteins. The arrows indicate the cleavage sites of the signal peptide in yeast and *E. coli*. Numbers on the amino acid residues indicate the positions relative to the amino terminus of the hybrid protein precursor.

THIS PAGE BLANK (USPTO)

shown in Fig. 8, after 60 min of labelling with [35 S]-methionine, all these mutant proteins were processed to the same degree as the non-mutant lipo- β -lactamase gene products. In addition, localization experiments with these mutant strains showed that the mature form of lipo- β -lactamase was found in the periplasm (data not shown). Thus, these signal sequence mutations do not appear to effect either processing or secretion of lipo- β -lactamase in yeast; however, there is still a possibility that there may be some differences in the kinetics of processing and translocation of these mutant proteins into the yeast periplasm. The altered mobility of the lipo- β -lactamase precursors results from the amino-acid alterations in the signal peptide.

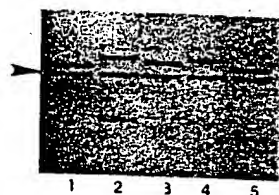


Fig. 8. Production and processing of lipo- β -lactamase. Cells harbouring pCOP1, pCOP2, pCOP3, pCOP4 and pCOP5 (lanes 1–5, respectively) were labelled with [35 S]-methionine for 60 min and subjected to SDS polyacrylamide gel electrophoresis, as described in Fig. 2. The arrow shows the position of processed lipo- β -lactamase.

Discussion

The secretion of proteins containing heterologous signal peptides has met with mixed success in yeast. Whereas a number of eukaryotic signal peptides are functional in yeast (Fujita *et al.*, 1986; Hitzeman *et al.*, 1983; Oberto and Davidson, 1985; Valenzuela *et al.*, 1982; Wood *et al.*, 1985), efficient secretion directed by prokaryotic signal sequences has not been reported in this organism (Schekman, 1985). In the present study, we have demonstrated that the prokaryotic hybrid protein lipo- β -lactamase is expressed in high amounts, and a significant portion of the protein synthesized is processed and secreted as an active enzyme into the yeast periplasm. Previous studies (Schekman, 1985; Roggenkamp *et al.*, 1983; Roggenkamp *et al.*, 1981) in which the native β -lactamase gene was expressed in yeast showed that the protein was presumably processed by non-specific proteases. Unlike this study, however, the processed form of β -lactamase accumulated in the yeast cytoplasm and was not translocated across the cytoplasmic membrane. Thus it appears that the signal sequence, plus nine amino acid residues from the amino terminus of the *E. coli* major lipoprotein, mediated successful translocation of lipo- β -lactamase into the yeast periplasm. In this respect, the lipoprotein signal sequence functions like some yeast signal sequences (MF α 1 and SUC 2) that can also direct the secretion of β -lactamase in yeast (Mileham *et al.*, 1986).

The cleavage of the lipo- β -lactamase signal peptide in yeast occurred between cysteine at position 21 and serine

22, which is one amino acid removed from the *E. coli* signal peptidase II processing site (see pCOP1 in Fig. 7). The cleavage of eukaryotic signal sequences after a cysteine residue has been demonstrated for a number of proteins such as immunoglobulin kappa and heavy chains, antithrombin III from humans, and ribulose biphosphate carboxylase from several plants (Watson, 1984). When the signal peptidase II processing site is altered, processing of lipo- β -lactamase in *E. coli* occurs via signal peptidase I at residue 25 (between an alanine and a lysine residue; Ghayeb *et al.*, 1985). In yeast, alteration of the signal peptidase II site, which is one residue away from the yeast cleavage site, does not block processing or secretion of lipo- β -lactamase (pCOP5 in Figs 7 and 8). It will be interesting to investigate whether alteration of either cysteine or serine of the yeast lipo- β -lactamase cleavage site will block secretion or processing. It is worth noting that alteration of the yeast invertase signal sequence cleavage site does slow down secretion (Schauer *et al.*, 1985).

The alteration of the basic domain of the lipoprotein signal sequence so that the net charge on that domain is -2 (from $+2$), severely inhibits processing and secretion of lipoprotein (Vlassuk *et al.*, 1983) and lipo- β -lactamase (Lunn and Inouye, 1987) in *E. coli*. Even though a net negative charge is not found in the corresponding domain of eukaryotic signal sequences, such an alteration in the lipo- β -lactamase signal sequence has no apparent effect on the level of processing or secretion of lipo- β -lactamase in yeast. It has been suggested that the basic domain of the signal peptide in prokaryotes interacts with the negatively charged membrane. In eukaryotes, this role may be fulfilled by additional components of the secretory pathway that have not been demonstrated in prokaryotes (e.g. the signal recognition particle). The behaviour, in yeast, of the amino-terminal lipo- β -lactamase mutation described above, is consistent with the report that various insertions into the amino terminus of the invertase signal sequence does not block secretion of invertase in yeast (Brown *et al.*, 1984).

Secretion and processing of lipo- β -lactamase (Lunn and Inouye, 1987), and of many other prokaryotic secretory proteins (Pollitt and Inouye, 1986), are very sensitive to alterations in the hydrophobic domain of the signal peptide. Two such mutants in lipo- β -lactamase were examined: the first contained a deletion of four hydrophobic amino acid residues from the hydrophobic domain, and the second contained a substitution of two residues that are strong α -helix breakers for two alanines. The first of these mutations severely inhibits processing and secretion of lipo- β -lactamase in *E. coli*, whereas the latter slightly enhances processing or secretion of lipo- β -lactamase in the prokaryotic system. Neither of these alterations appear to affect either the amount of processing or secretion of lipo- β -lactamase in yeast. The tolerance of deletions in the

THIS PAGE BLANK (USPTO)

yeast invertase signal peptide has been demonstrated recently. Deletions retaining eight out of the eleven contiguous hydrophobic amino acid residues of the hydrophobic domain were found to be fully active for secretion of this protein (Perlman *et al.*, 1986). Thus the remaining hydrophobic amino acids of the lipo- β -lactamase hydrophobic core may be sufficient for secretion in yeast.

Cysteine at position 21 of lipoprotein and lipo- β -lactamase is modified by lipid in *E. coli*. In yeast, however, cysteine 21 is removed upon processing of lipo- β -lactamase. It is interesting to note that a number of glycoproteins undergo fatty-acid acylation in the endoplasmic reticulum. This modification, however, is subsequently removed at some stage during (or after) transport of these proteins to the cell surface (Wen and Schlesinger, 1984). Since more than 60% of the lipo- β -lactamase in yeast is not processed, it will be interesting to test whether any lipid modification does occur.

Processing of lipo- β -lactamase is different in the yeast and *E. coli* systems. Whereas changes in the structure of the signal peptide of lipo- β -lactamase and invertase are tolerated in yeast secretion, *E. coli* secretion is very sensitive to changes in any one of the structural domains of the signal sequence. Consistent with this notion is a recent paper by Kaiser *et al.* (1987) which demonstrates that up to 20% of human 'random' sequences will at least partially function in secretion of invertase. Except for a tendency of these sequences to be hydrophobic, it appears that the specificity of the signal sequence structure required for recognition is low. Recently, we have also shown that the mutant lipo- β -lactamase from pCOP2 was processed in a eukaryotic cell-free system as efficiently as the wild-type lipo- β -lactamase (Garcia *et al.*, 1987). These differences observed in the secretion of lipo- β -lactamase in *E. coli* and yeast may reflect fundamental differences in the secretory machinery of prokaryotes and eukaryotes.

Experimental procedures

Strains plasmids and growth conditions

The *Saccharomyces cerevisiae* strain DMM1-15A (leu2, ura3, ade2, his5) and the plasmid YEp51 were kindly provided by Dr J. Broach (Broach *et al.*, 1983). The growth medium (SD) used in all the experiments contained 0.67% (w/v) nitrogen base without amino acids and 2% (w/v) of either dextrose or galactose and was supplemented with the appropriate amino acids. Cultures were pre-grown to stationary phase (12–24 h) in SD dextrose media at 30°C. These cultures were then inoculated into fresh SD media and supplemented with the appropriate carbon source. Lithium acetate DNA-mediated transformation was used for *S. cerevisiae* (Ito *et al.*, 1983). *E. coli* strain SB221 and standard cloning techniques have been described previously (Ching and Inouye, 1985).

Cell labelling and immunoprecipitation

S. cerevisiae DMM1-15A cells containing the appropriate plasmid were grown to 2×10^8 cells ml⁻¹, harvested, and resuspended in fresh SD media lacking methionine. After growing the cells for 30 min at 30°C, 100 μ Ci [³⁵S]-methionine was added to a 5-ml culture and the cells were incubated for 30–60 min. The labelled cells were collected by centrifugation, washed once with fresh SD media, and resuspended in 300 μ l of TE (10 mM Tris, pH 8, and 1 mM EDTA). Cells were broken by vortexing with glass beads for 1 min and centrifuged to recover a clear supernatant. This extract was immunoprecipitated with β -lactamase antiserum and formaldehyde-treated *Staphylococcus aureus* A Cowan cells (Kessler, 1975) and then analysed by SDS polyacrylamide gel electrophoresis as described (Ghrayeb and Inouye, 1984). Quantitation of film darkening observed on autoradiograms was carried out using a Shimadzu model CS-910 dual wave-length scanner.

Southern blot analysis

Yeast DNA was prepared as described by Sherman *et al.* (1983). DNA was transferred to nitrocellulose essentially according to Southern (Southern, 1979). A 1.5 kb *Xba*I, *Bam*HI fragment from plasmid pJG311, containing the lipo- β -lactamase gene, was nick-translated and used as the probe. Hybridization was performed in 3 \times SSC (1 \times SSC = 0.15 M NaCl, 0.015 M sodium citrate) 3 \times Denhart's solution (Denhardt, 1966) and 0.5% SDS at 65°C for 12–16 h. The filters were washed in 3 \times SSC, 0.5% SDS at 65°C for 2 h, dried and then autoradiographed.

Localization of prolipo- β -lactamase

The subcellular localization of prolipo- β -lactamase was determined by a modification of the method of Harrington and Douglas (1980). [³⁵S]-methionine-labelled yeast cells harbouring pCOP-1 were washed and resuspended to 2×10^8 cells ml⁻¹ in 0.5 ml of 1 M sorbitol, 50 mM Tris, pH 7.4, 2 mM dithiothreitol, 10 mM MgCl₂, 20 mM sodium azide, and 0.5 mM phenylmethylsulfonyl fluoride (PMSF). Zymolyase-20T (Sieckbagaku Kogyo, Co. Ltd; 30 μ l of a 2 mg ml⁻¹ stock) was added, and the suspension was incubated at 30°C for 30 min with gentle shaking. Formation of spheroplasts was monitored microscopically. The spheroplasts were harvested in a Sorval clinical centrifuge at 1000 rpm for 5 min at room temperature (the supernatant fluid constituted the periplasmic fraction) and resuspended in 0.5 ml of 50 mM Tris, pH 7.4, 3 mM EDTA, 0.5 mM PMSF. The cells were broken by sonication, then centrifuged at 300 000 \times g for 20 min using a Beckman Model TL-100 tabletop ultracentrifuge. The soluble fraction was removed, the membrane pellet resuspended in 0.5 ml of the same buffer and all fractions then subjected to immunoprecipitation and SDS polyacrylamide gel electrophoresis as described above.

Enzyme assays and other procedures

Detection of β -lactamase activity in SDS polyacrylamide gels was performed as previously described (Tai *et al.*, 1985). The method, in summary, includes renaturation of β -lactamase by washing of SDS gels in the presence of ampicillin. Whatman paper stained

THIS PAGE BLANK (USPTO)

with a starch-iodine solution and saturated with ampicillin is then placed on the gel. The appearance of white bands at the position of active β -lactamase is due to the ability of penicilloic acid (produced by the action of β -lactamase) to reduce I_2 and thus to decolorize the blue starch-iodine complex. Galactosidase and glyceraldehyde-3-phosphate dehydrogenase assays were performed according to published methods (Tschopp et al., 1984; Christensen and Cirillo, 1972).

Sequential Edman degradation was carried out as previously described (Nakamura and Inouye, 1979). The desired protein band was eluted from fixed and stained SDS polyacrylamide gels by gentle shaking for 24 h at 30°C, with 2 ml of 50 mM Tris-HCl, pH 8.0 containing 1% SDS. The eluted protein was then precipitated at 4°C with 2 volumes of acetone after the addition of 13 mg of crystallized bovine serum albumin. The precipitate was collected and washed 6 times with 1 ml of trifluoroacetic acid and applied to a JEOL Sequence Analyser model JAS-47K.

Acknowledgements

We thank Drs Paul March and Stephen Pollitt for critical reading of this manuscript. We also thank Ms Peggy Yazulla and Ms Rita Cuniff for their skilful secretarial assistance. This work was supported by grant GM19043 from NI General Medical Sciences and grant NP387M from the American Cancer Society to M.I., the Dr Chaim Weizmann Fellowship to O.P., and grant GM10362 from NI General Medical Sciences to C.A.L.

References

- Broach, J., Li, Y., Wu, L.C., and Jayaram, M. (1983) Vectors for high-level inducible expression of cloned genes in yeast. In: *Experimental Manipulation of Gene Expression*. Inouye, M. (ed.) New York: Academic Press, pp. 83–117.
- Brown, P.A., Halvorson, H.O., Raney, P., and Perlman, D. (1984) Conformational alterations in the proximal portion of the yeast invertase signal peptide do not block secretion. *Mol Gen Genet* 197: 351–357.
- Ching, G., and Inouye, M. (1985) Expression of *Proteus mirabilis* lipoprotein gene in *Escherichia coli*: existence of tandem promoters. *J Biol Chem* 261: 4600–4606.
- Christensen, M.S., and Cirillo, V.P. (1972) Yeast membrane vesicles: isolation and general characteristics. *J Bacteriol* 110: 1190–1205.
- Dennhardt, D.T. (1966) A membrane-filter technique for the detection of complementary DNA. *Biochem Biophys Res Commun* 23: 641–646.
- Fujita, N., Nelson, N., Fox, T.D., Claudio, T., Lindstrom, J., Riezman, H., and Hess, G.P. (1986) Biosynthesis of the *Torpedo californica* acetylcholine receptor α subunit in yeast. *Science* 231: 1284–1287.
- Garcia, P.D., Ghrayeb, J., Inouye, M., and Walter, P. (1987) Wild type and mutant signal peptides of *Escherichia coli* outer membrane lipoprotein interact with equal efficiency with mammalian signal recognition particle. *J Biol Chem* 262: 9463–9468.
- Ghrayeb, J., and Inouye, M. (1984) Nine amino acid residues at the amino terminal of lipoprotein are sufficient for its modification, processing and localization in the outer membrane of *Escherichia coli*. *J Biol Chem* 259: 463–467.
- Ghrayeb, J., Lunn, C.A., Inouye, S., and Inouye, M. (1985) An alternate pathway for the processing of the lipoprotein signal peptide in *Escherichia coli*. *J Biol Chem* 260: 10961–10965.
- Harrington, C.R., and Douglas, L.T. (1980) Inhibition and activation of mannan synthesis in *Saccharomyces cerevisiae* spheroplast lysates. *J Bacteriol* 142: 852–858.
- Hitzeman, R.A., Leung, D.W., Perry, L.J., Kohr, W.J., Levine, H.L., and Goeddel, D.V. (1983) Secretion of human interferons by yeast. *Science* 219: 620–625.
- Inouye, M., and Halegoua, S. (1980) Secretion and membrane localization of proteins in *Escherichia coli*. *Crit Rev Biochem* 7: 339–371.
- Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) Transformation of intact yeast cells treated with alkali cations. *J Bacteriol* 153: 163–168.
- Kaiser, C.A., Poess, D., Grisafi, P., and Bofstein, D. (1987) Many random sequences functionally replace the secretion signal sequence of yeast invertase. *Science* 235: 312–317.
- Kessler, S.W. (1975) Rapid isolation of antigens from cells with a *Staphylococcal* Protein A-antibody immunoabsorbent: parameters of the interaction of antibody-antigen complexes with Protein A. *J Immunol* 115: 1617–1624.
- Lunn, C.A., and Inouye, M. (1987) Effects of prolipoprotein signal peptide mutations on secretion of hybrid prolipo- β -lactamase in *Escherichia coli*. *J Biol* 8318–8324.
- Mileham, A.J., Wales, R., and Boyd, A. (1986) Secretion of bacterial β -lactamase from *Saccharomyces cerevisiae*: a comparison of several yeast signal sequences. *Yeast* 2: Abstr. S249.
- Nakamura, K., and Inouye, M. (1979) DNA sequence of the gene for the outer membrane lipoprotein of *E. coli*: an extremely AT-rich promoter. *Cell* 18: 1109–1117.
- Oberio, J., and Davison, J. (1985) Expression of chicken egg white lysozyme by *Saccharomyces cerevisiae*. *Gene* 40: 57–65.
- Perlman, D., Raney, P., and Halvorson, H.O. (1986) Mutations affecting the signal sequence alter synthesis and secretion of yeast invertase. *Proc Natl Acad Sci USA* 83: 5033–5037.
- Pollitt, S., and Inouye, M. (1986) Structure and functions of the signal peptide. In: *Bacterial Outer Membranes as Model Systems*. Inouye, M. (ed.) New York: John Wiley and Sons, Inc., pp. 117–140.
- Roggenkamp, R., Dargatz, H. and Hollenberg, C.P. (1985) Precursor of β -lactamase is enzymatically inactive: Accumulation of the preprotein in *Saccharomyces cerevisiae*. *J Biol Chem* 260: 1508–1512.
- Roggenkamp, R., Hoppe, J., and Hollenberg, C.P. (1983) Specific processing of the beta-lactamase precursor in *Saccharomyces cerevisiae*. *J Cell Biochem* 22: 141–149.
- Roggenkamp, R., Kustermann-Kuhn, B., and Hollenberg, C.P. (1981) Expression and processing of bacterial β -lactamase in the yeast *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 78: 4466–4470.
- Schauer, I., Emr, S., Gross, C., and Schekman, R. (1985) Invertase signal and mature sequence substitutions that delay inter-compartmental transport of active enzyme. *J Cell Biol* 100: 1664–1675.
- Schekman, R. (1985) Protein localization and membrane traffic. In: *Ann Rev Cell Biol*. Palade, G. (ed.) Vol. 1. Palo Alto, California: Annual Reviews Inc., pp. 115–143.
- Schekman, R. (1982) Protein export in yeast. In: *Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression*. Strathern, J.N., Jones, F.W., and Broach, J.R. (eds). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, pp. 651–652.
- Sherman, F., Fink, G.R., and Hicks, J.B. (1983) Isolation of yeast nuclear and mitochondrial DNA. *Methods in Yeast Genetics*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, pp. 77–80.

THIS PAGE BLANK (USPTO)

- Smith, R.A., Duncan, M.J., and Moir, D.T. (1985) Heterologous protein secretion from yeast. *Science* 229: 1219-1224.
- Southern, E.M. (1979) Gel electrophoresis of restriction fragments. *Meth Enzymol* 68: 152-176.
- Tai, P.C., Zyk, N., and Citri, N. (1985) *In situ* detection of β -lactamase activity in sodium dodecyl sulfate-polyacrylamide gels. *Anal Biochem* 144: 199-203.
- Tschopp, J., Esmon, P.A., and Schekman, R. (1984) Defective plasma membrane assembly in yeast secretory mutants. *J. Bacteriol* 160: 966-970.
- Valenzuela, P., Medina, A., Rutter, W.J., Ammerer, G., and Hall, B.D. (1982) Synthesis and assembly of hepatitis B virus surface antigen particles in yeast. *Nature* 298: 347-350.
- Vlasuk, G.P., Inouye, S., Ito, H., Itakura, K., and Inouye, M. (1983) Effects of the complete removal of basic amino acid residues from the signal peptide on secretion of lipoprotein in *Escherichia coli*. *J Biol Chem* 258: 7141-7148.
- Watson, M.E.E. (1984) Compilation of published signal sequences. *Nucl Acids Res* 12: 5145-5164.
- Wen, D., and Schlesinger, M.J. (1984) Fatty acid-acylated proteins in secretory mutants of *Saccharomyces cerevisiae*. *Mol Cell Biol* 4: 688-694.
- Wood, C.R., Boss, M.A., Kenten, J.H., Calvert, J.E., Roberts, N.A., and Emtage, J.S. (1985) The synthesis and *in vivo* assembly of functional antibodies in yeast. *Nature* 314: 446-449.

THIS PAGE BLANK (USPTO)

Importance of the Propeptide Sequence of Human Preproparathyroid Hormone for Signal Sequence Function*

(Received for publication, May 31, 1988)

Kristine M. Wirent†, John T. Potts, Jr., and Henry M. Kronenberg

From the Endocrine Unit, Massachusetts General Hospital and the Department of Medicine, Harvard Medical School, Boston, Massachusetts 02114

The function of amino-terminal pro-specific peptides (pr peptides), sequences often found on intermediate precursor forms of secreted proteins, is poorly understood. Human preproparathyroid hormone (prepro-PTH), a precursor protein containing such a propeptide, is initially synthesized as a precursor containing a 25-amino acid signal sequence, a 6-amino acid propeptide, and the 84-amino acid mature secreted peptide. Cloned cDNA encoding prepro-PTH and synthetic ligonucleotides were used to generate a mutant missing precisely the pro-specific sequences. The effects of this deletion on signal sequence function and on secretion *per se* were assessed after expression of the mutant cDNA in intact cells and in a cell-free translation system using synthetic mRNA in the presence of microsomal membranes. The mutant precursor protein was inefficiently translocated and cleaved, and cleavage occurred both at the normal site and within the signal sequence. Thus, for the eukaryotic protein prepro-PTH, sequences immediately downstream and separate from the classically defined signal sequence facilitate accurate and efficient signal function.

Secretory proteins are generally translated with amino-terminal sequences, termed signal sequences, that mediate interaction with the membrane of the rough endoplasmic reticulum (Blobel and Dobberstein, 1975a, 1975b; Hortsch and Meyer, 1986; Walter *et al.*, 1984). In some cases, after signal cleavage the resultant protein is identical to the mature secreted form. Often, however, intermediate precursors, called pro-proteins, missing the signal sequence but containing peptide residues not found in the mature secreted protein, are found intracellularly. The extra, pro-protein-specific regions, termed propeptides, are proteolytically processed from the intermediate precursor only at the distal aspect of the secretory pathway, in the trans-Golgi network (Giffiths and Simons, 1986) and developing secretory granule (Orci *et al.*, 1986). Cleavage occurs after dibasic residues, by enzymes still being characterized (Steiner *et al.*, 1984).

Unlike signal sequences, pro-specific sequences are unlikely to have one uniform function from protein to protein, because their lengths, locations in the precursor, and structures vary so greatly (Docherty and Steiner, 1982). Possible functions of

propeptides include: 1) ensuring correct folding of the protein (Steiner and Clark, 1968); 2) providing a minimum length needed for transport across the endoplasmic reticulum; 3) providing linkage for coordinate translation of related secreted peptides, such as the cleavage products of proopiometanocortin (Herbert and Uhler, 1982); 4) conferring a diverse spectrum of actions to a family of alternatively cleaved precursors, such as the precursor of somatostatins SRIF-14 and SRIF-28 (Noe and Spiess, 1983); 5) acting as sorting sequences (Guan and Rose, 1984) to guide precursors through the secretory pathway from the rough endoplasmic reticulum to the Golgi, where cleavage of the propeptide ensues; 6) directing covalent modification of protein, as suggested by the observation that the propeptide plays a role in post-translational γ -carboxylation of Factor IX (Jorgensen *et al.*, 1987); or 7) forming a functional unit with the signal peptide, to assure optimal signal function.

The site of cleavage of the signal from any precursor protein can be predicted fairly reliably, with rules based on the observed frequency of certain amino acid residues, relative to the cleavage site, in already characterized signals (Perlman and Halvorson, 1983; von Heinje, 1983, 1984, 1986a, 1986b). From such analyses, a pattern of acceptable and forbidden sequences bordering the cleavage site emerged. Acceptable cleavage domains conform to the following rules: the residue at position -1 from the cleavage site must be small (*i.e.* Ala, Ser, Gly, Cys, Thr, or Gln); and the residue at position -3 must not be aromatic, charged, large and polar, or a proline. This (-3, -1) rule, combined with the expected distribution of other amino acids within the cleavage domain (-13 to +2), has been used to construct a weight matrix to calculate the probability of cleavage at a specific site within the cleavage domain of eukaryotic proteins. The analysis has shown further that signal cleavage usually occurs within a window 4-10 residues from the end of the hydrophobic core and that sequences immediately distal to the cleavage site tend to minimize choices for alternate cleavage sites (von Heinje, 1986b).

To investigate the function of propeptides, we have chosen as a model the secretory protein preproparathyroid hormone (prepro-PTH).¹ Biosynthetic studies indicate that, in the parathyroid gland, parathyroid hormone (PTH) is formed first as the larger peptide prepro-PTH. This precursor undergoes two successive proteolytic cleavages to yield PTH; the transitory amino-terminal signal peptide is cleaved off in the rough endoplasmic reticulum co-translationally to yield the intermediate precursor pro-PTH (Cohn and Elting, 1983; Habener *et al.*, 1979) which is processed later in the Golgi to produce PTH (Habener *et al.*, 1979). Secreted mature PTH is

* This investigation was supported in part by National Institutes of Health Grant AM27758. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported by National Research Service Award DK07708 and National Institutes of Health Training Grant AM07028. Current address: Geriatric Research, Education and Clinical Ctr., American Lake Veterans Administration Medical Ctr., Tacoma, WA 98493.

¹ The abbreviations used are: PTH, parathyroid hormone; SDS, sodium dodecyl sulfate; *gpt*, guanine-xanthine phosphoribosyl transferase gene.

THIS PAGE BLANK (USPTO)

an unglycosylated 84-residue protein whose classical actions are to raise blood calcium and lower blood phosphorus by acting on bone, kidney, and, indirectly, on intestine (Habener *et al.*, 1984). The primary translation product, prepro-PTH, has a 25-amino acid signal, followed by the highly basic 6-residue propeptide sequence.

The small size of prepro-PTH's propeptide makes unlikely most of the hypotheses about function outlined above. The human propeptide (Lys-Ser-Val-Lys-Lys-Arg), is cleaved from pro-PTH just before secretion, presumably in the trans-Golgi tubular network. Processing occurs after the dibasic residues Lys-Arg. Pro-PTH is not secreted from cells (Habener, 1979), and neither the pro-specific fragment nor any of the its possible degradation products accumulates in the cell (Habener, 1979); thus, any role for the propeptide must be an intracellular one.

The propeptide of pro-PTH should serve as an appropriate model for analyzing the function of other short amino-terminal propeptide sequences. To test for potential functions, we have used a human prepro-PTH cDNA clone to generate a mutant gene (called pre(Δ pro)-PTH) lacking precisely and only the pro-specific sequences. The propeptide deletion was created by oligonucleotide-directed mutagenesis. Two independent but parallel approaches were used in the analysis: the mutant pre-PTH cDNA was integrated stably into eukaryotic secretory cells using a retroviral vector; furthermore, the cDNA was expressed by translation of synthetic mRNA in a cell-free extract containing microsomal membranes. The studies reported here define a role for the propeptide in signal function, since the pre(Δ pro)-PTH mutant precursor is 1) inefficiently translocated, 2) inefficiently processed to PTH-related peptides, and 3) cleaved at both the normal site and at an alternate site located two amino acids proximal to the normal site within the signal sequence. These results indicate that, in the eukaryotic protein prepro-PTH, sequences immediately downstream and separate from the classically defined signal facilitate several aspects of signal sequence function.

EXPERIMENTAL PROCEDURES

Materials

Restriction enzymes, T4 DNA ligase, and *Escherichia coli* RNA polymerase were from New England BioLabs. Formalin-fixed *Staphylococcus aureus* (IgG-sorb) was from the Enzyme Center, Inc. L-[³⁵S] Methionine and L-[³H]leucine were from Amersham Corp. Trypsin and α -chymotrypsin were from Sigma. RQ1 DNase and canine pancreatic microsomal membranes were from Promega Biotec.

Methods

Cell Culture—GH₄C₁ rat pituitary cells were grown in Costar plasticware in Dulbecco's modified Eagle's medium with high glucose (4.5 g/liter) and 10% fetal calf serum (Gibco), in 95% air and 5% CO₂. The cells were replated every 5–7 days, using 0.05% trypsin and 0.02% EDTA to disperse the cells. The ψ -2 cell line (Mann *et al.*, 1983), a clonal NIH 3T3 cell line stably transfected with modified Moloney murine leukemia virus, was a gift of R. Mulligan (Massachusetts Institute of Technology).

In Vitro Transcription and Translation—pGL101-based (Guarente *et al.*, 1980) plasmids containing the UV5 *lac* promoter region were transcribed *in vitro* by *E. coli* RNA polymerase as described (Kronenberg *et al.*, 1983). The transcription reaction was subjected to RQ1 DNase digestion (to remove the DNA template) and then extracted with phenolchloroform and ethanol-precipitated. The mRNA concentration was measured spectrophotometrically. The resulting mRNA was translated in a nuclease-digested reticulocyte lysate in the presence of either [³⁵S]methionine or [³H]leucine, as described previously (Kronenberg *et al.*, 1983). To test for membrane translocation and signal peptidase cleavage, rough microsomes from canine pancreas were included in the translation reaction, either prepared by the

method of Shields and Blobel (1978), or purchased from Promega Biotec.

To then assay for cotranslational translocation, trypsin, α -chymotrypsin, and, in some reaction, Triton X-100 (1% final concentration) as a control were added to the cell-free system 30 min after the initiation of translation. As an additional control, microsomal membranes were added to some reactions post-translationally. Reactions were then incubated 90 min on ice, followed by the addition of aprotinin (150 KIU) to all reactions, and then Triton X-100 (0.2% final concentration) to those without prior Triton X-100 addition. Radiolabeled proteins were then subjected to immunoprecipitation.

Immunoprecipitation and Gel Electrophoresis—Reactions containing radiolabeled proteins were immunoprecipitated, after preclearing overnight, as described (Hellerman *et al.*, 1984). The pellets were extensively washed in RIPA buffers (10 mM Tris-HCl, pH 7.2, 1% Triton X-100, 0.1% SDS, 1% deoxycholate, 5 mM EDTA, and 0.02% NaN₃), dissociated in sample buffer, and run on 15–20% SDS-polyacrylamide gradient slab gels (Laemmli, 1970). Gels were incubated in En³hance (Du Pont-New England Nuclear) according to directions supplied by the manufacturer and dried. Proteins were visualized with fluorography on Kodak XAR-S film. Quantitation of bands on the gels was performed by scanning densitometry on an LKB Ultrascan XL laser densitometer.

Amino-terminal Sequence Determination—PTH-peptides were radiolabeled in the presence of [³⁵S]methionine either in intact cells or in reticulocyte lysates programmed with mRNA from the *in vitro* transcription reactions. Peptides were isolated by immunoprecipitation as described above, followed by SDS-polyacrylamide gel electrophoresis on high resolution (27 cm) slab gels. Gels were dried without En³hance, autoradiography was performed, and the region of the gel containing PTH peptides was excised with a razor blade. The eluate from overnight incubation of the gel slice with a solution containing 1% triethylamine, 0.1% SDS, and 200 μ g/ml gelatin was lyophilized and subjected to automated Edman degradation using a Beckman Model 890C protein sequencer and a 0.1 M Quadrol program as described previously (Kronenberg *et al.*, 1983).

Plasmid Constructions—The methods for preparation of plasmid DNA, cleavage with restriction enzymes, purification of fragments, and ligation reactions were performed as described (Hendy *et al.*, 1981). A plasmid encoding pre(Δ pro)-PTH was constructed using oligonucleotide-directed mutagenesis (see "Results"). The synthetic 82-base oligonucleotide fragments were synthesized on a Model 380A DNA synthesizer from Applied Biosystems, by the solid-phase phosphoramidite method. Sequencing of the mutant cDNA to confirm the sequence (after cloning), was by the chemical method of Maxam and Gilbert (Maxam and Gilbert, 1980). The plasmid pZIP-v-gpt was a gift of R. Mulligan.

DNA Transfections and Viruses—DNA transfections were performed with 10 μ g of plasmid DNA by the procedure of Graham and Van der Eb (1973), as modified by Parker and Stark (1979). Virus infection was performed in the presence of 8 μ g of Polybrene/ml for 2.5 h, as described (Hellerman *et al.*, 1984). Aminopterin was omitted from the gpt-selection medium for GH₄C₁ cells (Mulligan and Berg, 1981).

To radiolabel cells for pulse-chase studies, confluent 100-mm dishes were rinsed twice in methionine-deficient Dulbecco's modified Eagle's medium containing 10% dialyzed fetal calf serum. Cells were then incubated with 0.5 mCi of [³⁵S]methionine in 3 ml of medium. Chase was initiated by rinsing cell layers once with phosphate-buffered saline and refeeding with fresh normal Dulbecco's modified Eagle's medium containing 30 mg/l of nonradioactive methionine. Incubations were terminated by lysis in detergent-containing RIPA buffer with 500 μ g/ml phenylmethylsulfonyl fluoride added. Cell lysates were forced repeatedly through a 22-gauge needle to shear DNA, and both media and cell lysates were spun for 15 min at 15,000 \times g. Samples were then processed for immunoprecipitation and SDS-polyacrylamide gel electrophoresis as described above.

RESULTS

Construction of pre(Δ pro)-PTH Deletion Mutant cDNA—Oligonucleotide-directed mutagenesis was used to delete the 18 nucleotides that encode the 6-amino acid propeptide from prepro-PTH. Synthetic HindIII and BamHI cleavage sites had previously been inserted at the ends of human prepro-PTH cDNA (Born *et al.*, 1986). Employing a BglIII site fortuitously located precisely at the end of the prepro-sequence,

THIS PAGE BLANK (USPTO)

the entire "prepro" coding sequence was excised by digestion with *Hind*III and with *Bgl*II. Two 82-base oligonucleotides were synthesized and provided the complementary strands encoding the entire 25-amino acid prepeptide with appropriate cohesive ends. The oligonucleotides were annealed by slow cooling from 68 °C to room temperature, in low salt buffer (10 mM Tris, pH 7.5, 10 mM MgCl₂, 1 mM dithiothreitol) at 100 ng/μl for each strand. This double-stranded sequence was then ligated just upstream from the coding sequences for mature PTH. This construction thus deleted precisely and only the sequences encoding the propeptide, generating pre(Δpro)-PTH cDNA. The sequence, orientation and reading frame of the deletion plasmid were confirmed by DNA sequencing using the chemical method (Maxam and Gilbert, 1980).

From the parent construct, two other plasmids were made; one based upon an *E. coli* expression plasmid pGL101 (Guarente *et al.*, 1980) for studies employing cell-free translation assays using synthetic mRNA, and one based upon the plasmid pZIP-v-gpt (analogous to pZIP-neoSV(X)1 (Cepko *et al.*, 1984) but containing the *gpt* gene) used to generate recombinant retrovirus that, in turn, was used to infect GH₄C₁ cell lines. In the pGL101-based construct, transcription of pre(Δpro)-PTH mRNA was driven from the *lac* promoter *in vitro* using *E. coli* RNA polymerase. Messenger RNA produced in this reaction was then used to program cell-free translation of pre(Δpro)-PTH protein. Clonal GH₄C₁ cell lines stably expressing mutant human pre(Δpro)-PTH cDNA were obtained by infection of cultures with recombinant retrovirus containing the mutant cDNA, using transient rescue, as follows: ψ-2 cells, which contain a defective Moloney murine leukemia provirus (Mann *et al.*, 1983) were transfected with the retroviral construct by calcium-phosphate precipitation. These cultures package infective viral particles that contain only the pre(Δpro)-PTH RNA. Medium harvested from these cultures 20 h later contained the recombinant retrovirus that was used to infect GH₄C₁ cells. Infected cultures were grown under guanine-xanthine phosphoribosyl transferase (*gpt*) selection in the presence of mycophenolic acid. Clones expressing *gpt* activity, and presumably pre(Δpro)-PTH, were isolated.

Processing of Pre(Δpro)-PTH in Cell-free Extracts—Translation reactions using synthetic mRNA, carried out in reticulocyte lysate in the presence of canine pancreatic microsomal membranes, were used to assay *in vitro* the early events in the processing of pre(Δpro)-PTH. In this assay, proteins are cotranslationally translocated across the membrane, and signal sequences are cleaved by signal peptidase, located on the inner face of the membrane (Jackson and Blobel, 1977; Kreil, 1981). The mRNA was transcribed *in vitro* from normal and mutant linearized plasmids (20 μg), using *E. coli* RNA polymerase. The template DNA was digested with RQ1 DNase (20 units), and the mRNA was extracted with phenolchloroform and ethanol-precipitated. Messenger RNA concentration was measured spectrophotometrically after resuspending the pellets in H₂O. Both normal and mutant constructs yielded identical amounts of message (data not shown). Synthesis of prepro-PTH was performed with undiluted mRNA while synthesis of pre(Δpro)-PTH was performed with mRNA diluted 1:5 and 1:10 (see below).

Translation reactions were performed in the absence or in the presence of increasing concentrations of rough microsomes from canine pancreas. To assay for proteolytic processing by signal peptidase, proteins were translated in the presence of [³H]methionine, immunoprecipitated with affinity-purified anti-PTH antisera, separated on 15–20%

SDS-polyacrylamide gradient slab gels, and processed for fluorography.

The results of translation reactions directed by pPTHlac3 (Born *et al.*, 1986), a pGL101-based plasmid encoding normal human prepro-PTH, are shown in the upper panel of Fig. 1. The primary translation product in the absence of membranes (0 equivalents) was prepro-PTH. The more rapidly migrating immunoprecipitable proteins are PTH-related peptides that are generated by internal initiation at the methionines at positions 7, 39, and 49 in the prepro-PTH molecule. The identity of all PTH-related peptides was established by amino-terminal radiosequence analysis (data not shown). Translation in the presence of increasing concentration of rough microsomes from canine pancreas (1, 2, and 4 equivalents) resulted in cleavage of the signal, producing pro-PTH. Increasing the concentration of membranes generated more processed pro-PTH and less prepro-PTH, with almost complete conversion at the highest dose (4 equivalents) of membranes employed.

When message concentrations were equal, more pre(Δpro)-PTH peptide was translated than normal prepro-PTH (data not shown). This variable efficiency of translation presumably results from the differences in the mRNA structures: the prosequence deletion and differences in the 5'- or 3'-untranslated regions.

Relatively equivalent amounts of protein were synthesized when the mutant mRNA was diluted 1:5 relative to the normal mRNA (Fig. 1). Even under conditions in which less mutant

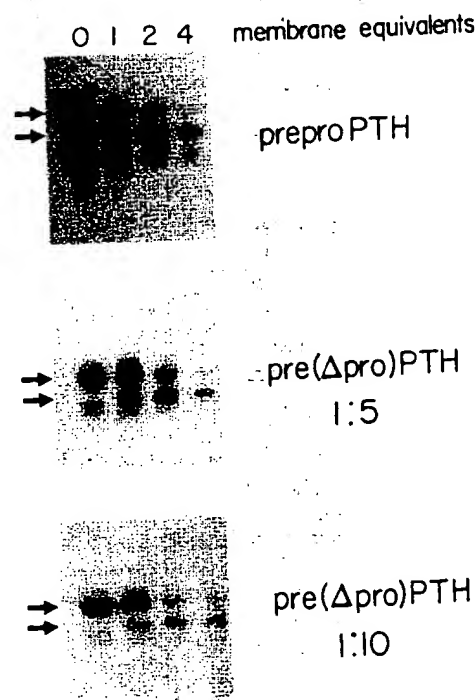


FIG. 1. Cleavage of PTH-related precursors by increasing doses of microsomal membranes. Immunoprecipitates of PTH-related peptides after translation with adjusted mRNA concentrations in the presence of indicated equivalents of microsomal membranes added cotranslationally. See text for explanation of bands. Arrows indicate precursor and processed products. The percent conversion of precursor to product was calculated from the integrated peak area for precursor and processed forms: prepro-PTH at 1 eq = 45%, 2 eq = 71%, and 4 eq = 90%; pre(Δpro)-PTH 1:5 at 1 eq = 40%, 2 eq = 56%, and 4 eq = 76%; pre(Δpro)-PTH 1:10 at 1 eq = 30%, 2 eq = 51%, and 4 eq = 66%. Results were adjusted to normalize for the 6 methionines in each processed product. The difference between normal prepro-PTH and pre(Δpro)-PTH is statistically significant at the $p = 0.037$ level.

THIS PAGE BLANK (USPTO)

peptide was synthesized (1:10 dilution), the processing was somewhat inefficient. Fig. 1 compares quantitatively the cleavage of normal prepro-PTH to the cleavage of pre(Δ pro)-PTH. The relevant bands were scanned, and the fraction of synthesized protein cleaved at each membrane concentration was calculated. Even at low protein concentrations, pre(Δ pro)-PTH peptide was inefficiently processed when compared to normal prepro-PTH. The difference between the processing of pre(Δ pro)-PTH and prepro-PTH was statistically significant at the $p = 0.037$ level using the nonparametric triangle test (Hollander and Wolfe, 1973).

Deletion of the propeptide from the prepro-PTH molecule therefore resulted in inefficient cleavage of the signal in a cell-free system, even though the actual signal sequence was unaltered in the mutant precursor. As will be discussed in more detail later, amino-terminal sequence analysis was carried out on the processed PTH product to determine the site of cleavage. The predominant processed protein produced in the reticulocyte lysate was authentic PTH, generated by cleavage of the signal sequence at the normal position (data not shown).

Translocation of Pre(Δ pro)-PTH through the Membrane—A proteolytic protection protocol was used to determine the location of precursor and product in the cell-free system, relative to the phospholipid bilayer. Since signal peptidase is located on the inner face of the membrane, any processed product must have been translocated through the membrane in order to have been a substrate for signal peptidase activity. However, translocation can occur without signal cleavage (Hortin and Boime, 1981), and consequently must be measured by an assay independent of cleavage. Limited trypsin/chymotrypsin digestion in the presence of membranes therefore more directly assesses translocation through the membrane. Trypsin and chymotrypsin were added after the completion of translation in the presence of rough microsomes from canine pancreas and [3 H]leucine, and samples were incubated for 90 min on ice. Radiolabeled proteins were immunoprecipitated after inactivating the proteases with aprotinin and subjected to SDS-polyacrylamide gel electrophoresis as described above.

Radiolabeled protein resistant to degradation is presumed to be protected within a membrane-bound space, isolated from protease digestion. In the presence of rough microsomes from canine pancreas, pro-PTH cleaved from prepro-PTH was completely protected from limited digestion (Fig. 2, lanes 3 and 6), while precursor prepro-PTH was sensitive to digestion and disappeared. Similarly, processed PTH from pre(Δ pro)-PTH was also protected (lanes 11 and 14), while the precursor

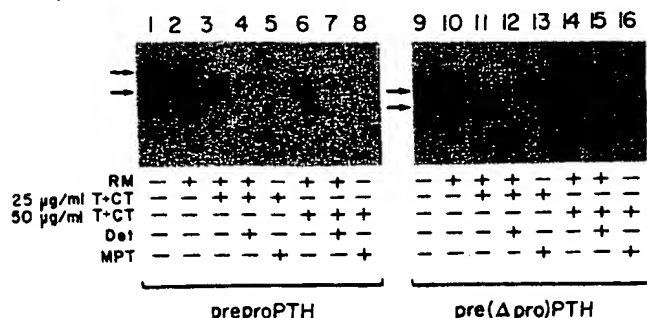


FIG. 2. Translocation of PTH-related peptides. Translation reactions were performed in the presence of rough microsomes (RM) in indicated lanes; membranes were added only post-translationally (MPT) in indicated lanes. After translation, trypsin and chymotrypsin (T + CT) were added at the concentrations indicated. Triton X-100 was added at the same time as the protease in the indicated lanes (Det). Precursor and processed products are indicated by arrows.

was not. The first two lanes in each set again demonstrate the inefficient processing of pre(Δ pro)-PTH seen in Fig. 1 (compare lane 2 to lane 10 in Fig. 2). The translocation of pre(Δ pro)-PTH was therefore less efficient than the translocation of normal prepro-PTH. The PTH-peptides were inherently digestible, since disruption of membrane integrity with detergent led to digestion of both the precursor and the product (lanes 4, 7, 12, and 15) radiolabeled proteins at two different protease concentrations. The peptides were not protected from digestion by a nonspecific hydrophobic interaction with membranes, since addition of microsomes after translation did not protect them from digestion (lanes 5, 8, 13, and 16).

Processing and Secretion of Pre(Δ pro)-PTH by Cells Infected with Recombinant Retrovirus—The cell-free studies outlined above indicate that processing of pre(Δ pro)-PTH is inefficient. Cell-free systems best assay the early steps in secretion. In order to examine the later stages of secretion, and to confirm the cell-free data, clonal secretory cell lines expressing pre(Δ pro)-PTH were established using recombinant retrovirus. The rat pituitary GH $_4$ C $_1$ tumor cell line was chosen as the recipient cell line for the transfection studies, because it is a well characterized prolactin-secreting line (Tashjian *et al.*, 1978). These cells do contain some vesicles and secretory granules at the cell periphery, and respond in a physiological manner to thyrotropin-releasing hormone by increased release of prolactin (Tashjian *et al.*, 1978). PTH-derived peptides use the normal secretory pathway in these cells, since clones expressing prepro-PTH respond to a thyrotropin-releasing hormone challenge by increased release of PTH into the media (Hellerman *et al.*, 1984). The pre(Δ pro)-PTH recipient GH $_4$ C $_1$ clones derived by *gpt* selection were characterized by immunoprecipitation using affinity-purified goat anti-human PTH antisera. Four independent clones were established and shown to secrete identical products; one was used for the detailed characterization of pre(Δ pro)-PTH processing.

Cells were pulse-labeled for 15 min with [35 S]methionine and then chased for the times indicated (Fig. 3a) with medium containing excess nonradioactive methionine. Media and cell lysate samples were immunoprecipitated and separated on

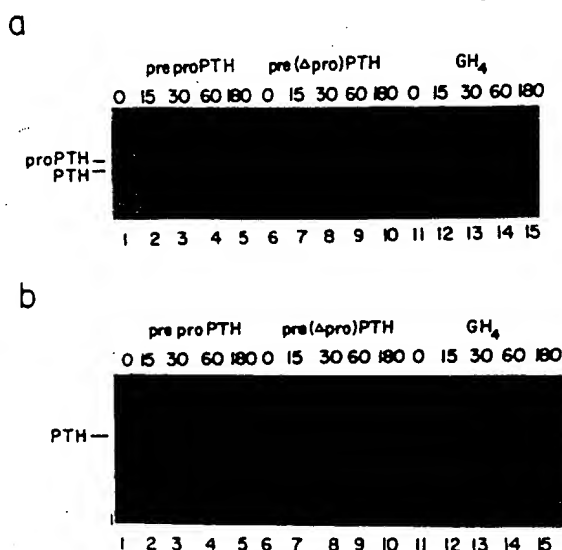


FIG. 3. Pulse-chase analysis of control GH $_4$ C $_1$ cells and GH $_4$ C $_1$ cells synthesizing prepro-PTH pre(Δ pro)-PTH. Immunoprecipitates of cell lysates (a) and media (b) were electrophoresed after cells were labeled with [35 S]methionine for 15 min and chased with nonradioactive methionine for the indicated times.

THIS PAGE BLANK (USPTO)

15–20% SDS-polyacrylamide gradient slab gels and then processed for fluorography. The results of a pulse-chase analysis of a clone expressing normal prepro-PTH are shown in Fig. 3a. As in normal parathyroid cells, no prepro-PTH precursor was detected after a 15-min pulse labeling of such cells (lane 1), because cleavage and removal of the signal sequence is presumably cotranslational. The predominant products immunoprecipitated at the earliest time point were pro-PTH and PTH (lane 1). With time, pro-PTH was converted to PTH (lanes 2–5), and PTH was found secreted into the medium at 15–30 min post-chase (Fig. 3b). Neither precursor prepro-PTH nor pro-PTH were secreted into the media from these cultures.

In cells that had integrated the pre(Δ pro)-PTH recombinant retrovirus, the protein that co-migrated with the uncleaved precursor pre(Δ pro)-PTH marker protein produced in a cell-free extract without membranes (marker not shown) was the predominant PTH-related protein at the end of the 15-min pulse (Fig. 3a). The mutant clonal cell line contains less peptide than the clonal line expressing normal prepro-PTH, presumably because the site of integration affects the level of gene expression. The inefficient processing of the pre(Δ pro)-PTH precursor (lane 6) was in striking contrast to the efficient processing of normal prepro-PTH (which is therefore not visible in lane 1). The precursor pre(Δ pro)-PTH, although a predominant band at the earliest time point, rapidly disappeared from the cells. Since the precursor was not secreted into the media (Fig. 3b), it was either rapidly degraded or cleaved to the PTH-sized product in Fig. 3a.

In the medium from cultures expressing pre(Δ pro)-PTH, PTH-sized peptide was secreted with kinetics similar to that of native PTH, appearing around 30 min (Fig. 3b, lane 8). No bands that co-migrate with PTH-peptides were immunoprecipitated from untransfected GH₄C₁ (lanes 11–15).

Sequence Analysis of Products Secreted into the Medium— The sites of cleavage of normal prepro-PTH are accurate in GH₄C₁ cells (Hellerman *et al.*, 1984). To characterize the site of cleavage of the precursor pre(Δ pro)-PTH that generates the PTH-sized protein in GH₄C₁ cells, amino-terminal radiosequence analysis was performed. [³⁵S]Methionine was used to label the protein, since mature human PTH contains methionine residues at positions 8 and 18. The cleavage products secreted into the medium during a 2-h labeling of cells were eluted from a gel slice after preparative high resolution polyacrylamide gel electrophoresis and subjected to repetitive automated Edman degradation. The longer gel allowed the resolution of two closely spaced bands of nearly equal intensity. As can be seen in Fig. 4b, sequence analysis of the lower band from the medium revealed [³⁵S]methionine peaks at the expected positions of 8 and 18, indicating cleavage at the normal end of the signal sequence, after a glycine residue. This finding confirms the result found with *in vitro* processing (data not shown). However, sequence analysis of the upper band showed [³⁵S]methionine peaks at positions 10 and 20 (Fig. 4a). Cleavage to produce this peptide therefore occurred after serine 23 of the pre(Δ pro)-PTH precursor, 2 amino acids back into the signal sequence, yielding Asp-Gly-PTH as the processed product (Fig. 4c). Cleavage at this site conforms to the (–3, –1) rule, although this site is only three amino acid residues downstream from the end of the hydrophobic core of the signal. In the pulse-chase study (Fig. 3a) these two peptides were found both in the media and in the cell lysates. Both normal PTH and Asp-Gly-PTH appeared in both the cell lysate and media samples with the same kinetics. Removal of the propeptide therefore led to ambiguity in the site selected by signal peptidase for cleavage of the

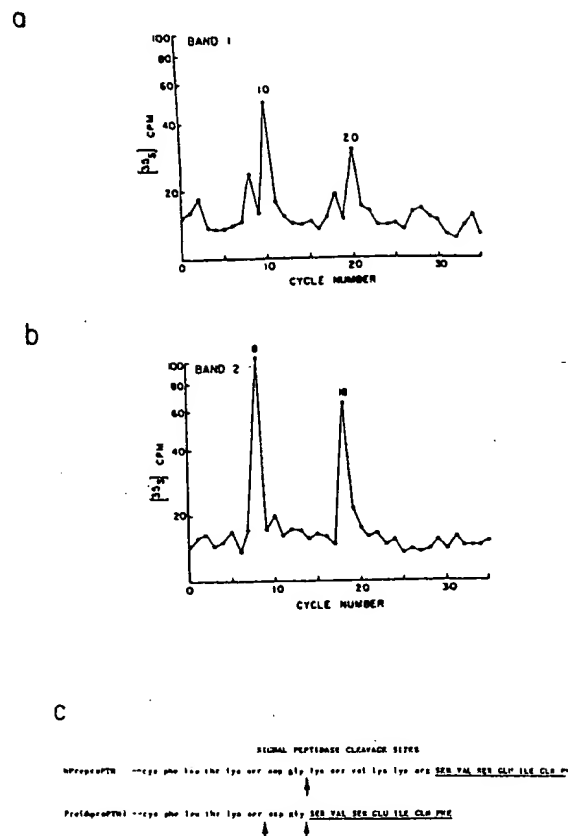


FIG. 4. Amino-terminal radiosequence analysis of proteins secreted from GH₄C₁ cells synthesizing pre(Δ pro)-PTH. Proteins were separated by gel electrophoresis after labeling of cells with [³⁵S]methionine. The sequence of the upper band (a) indicates that the methionines found at positions 8 and 18 of PTH are located at positions 10 and 20 in this protein. The sequence of the lower band (b) is that of PTH. The cleavage sites that are suggested by the methionine peaks are indicated in c.

signal peptide, but did not dramatically alter the kinetics of secretion. A second cleaved product was not detected in the cell-free translation reactions, in contrast (Figs. 1 and 2). However, sequence analysis of the product synthesized *in vitro* suggested that a small amount of Asp-Gly-PTH might be synthesized in the cell-free system (data not shown).

DISCUSSION

The role played by amino-terminal propeptides in the complex process of secretion has not been clarified. In order to characterize potential functions of propeptides, we have used human prepro-PTH as a model secretory protein, and deleted the hexapeptide prosegment by oligonucleotide-directed mutagenesis of cloned cDNA. This precise deletion created a cDNA encoding a peptide consisting of the naturally occurring 25-residue-long signal peptide from prepro-PTH fused in phase with the mature PTH molecule. We have used this cDNA to create vectors that program both *in vitro* reactions and intact cells for the expression of the pre(Δ pro)-PTH precursor peptide.

We have demonstrated that precise deletion of the propeptide-specific domain from the prepro-PTH molecule leads to inefficient translocation and processing of the signal sequence from the precursor pre(Δ pro)-PTH. Furthermore, the deletion of the pro-specific peptide leads to ambiguity in the signal cleavage sites chosen by signal peptidase. These results suggest that sequences immediately downstream from (carboxyl-terminal to) the classically defined signal sequence are im-

THIS PAGE BLANK (USPTO)

portant for proper signal function. The possible generality of this observation is suggested by the exonic organization of the human, bovine, and rat PTH genes (Vasicek *et al.*, 1983; Weaver *et al.*, 1984; Heinrich *et al.*, 1984). The first exon of the PTH genes contain 5'-noncoding sequences. Exon II contains both the entire signal and sequences encoding the first four amino acid residues of the propeptide domain (Vasicek *et al.*, 1983; Weaver *et al.*, 1984; Heinrich *et al.*, 1984). Given the hypothesis that exons are organized into functional domains (Gilbert, 1978), this organization supports the hypothesis that the propeptide segment, necessary for accurate cleavage and efficient translocation, be considered as part of the domain of the signal sequence. Many other signal-encoding exons similarly include portions of pro-encoding sequences.

Deletion of the propeptide portion of prepro-PTH does not dramatically alter the kinetics of secretion of PTH-related peptides in GH₄C₁ clones expressing pre(Δ pro)-PTH. Thus, the processed PTH-sized doublet travels the secretory pathway with relatively the same kinetics as does normal pro-PTH. The rapid secretion of PTH peptides from the cells synthesizing pre(Δ pro)-PTH suggests that the pro-specific sequence is not required for efficient transport of the PTH molecule from the endoplasmic reticulum to the Golgi.

Translocation of pre(Δ pro)-PTH through the membrane was less efficient than the translocation of wild type prepro-PTH in cell-free extracts. This result is particularly interesting and somewhat puzzling because the signal is unaltered in the pre(Δ pro)-PTH mutant peptide. The explanation for this result is not obvious. Perhaps the efficiency of cleavage by signal peptidase is reduced; in that case, the uncleaved precursor might fail to cross the membrane or even slip back into the cell-free equivalent of the cytoplasm. Alternatively, the propeptide may directly influence translocation by a mechanism unrelated to inefficient signal cleavage.

One of the most striking alterations in phenotype of the pre(Δ pro)-PTH precursor is the use of alternate cleavage sites by signal peptidase. In an attempt to define the features present in the mutant precursor peptide that might be responsible for the use of the alternate cleavage site, the signal sequence was characterized by von Heinje's probabilistic method (von Heinje, 1986b). This method allows comparison of the mutant precursor to sequences of other characterized precursor proteins to predict appropriate cleavage sites. Since the residues bordering the alternative cleavage site are identical in the mutant and normal peptide, there is no difference in the calculated probability that the alternate site, within the signal sequence, will be used. While the probability of cleavage at the alternate site is high for both peptides (data not shown), the alternate site is, in fact, only used in the context of the pre(Δ pro)-PTH peptide. This site is used even though it is only 3 residues from the end of the hydrophobic core, outside of the cleavage window. Furthermore, a second site with high probability, after the serine at position 3 in mature PTH, is within the window for cleavage, and is used in related mutants missing the last several residues of the signal sequence as well as the propeptide,² but not in the pre(Δ pro)-PTH precursor.

Examination of the primary structure of the pre(Δ pro)-PTH mutant thus fails to completely predict its phenotype. It has been suggested (Perlman and Halvorson, 1983) that an alteration of β -turn potential and interruption of helical domains may be important structural features in determining the site of signal peptidase cleavage. Secondary structure

predictions of the prepro-PTH molecule (Chou and Fasman, 1978) demonstrate that the pro-specific segment has no marked tendency to form secondary structure, either of the helix or β -sheet type (data not shown). This possibly unstructured region may be a preferred substrate for signal peptidase and/or may provide flexibility needed for efficient translocation. Deletion of the propeptide juxtaposes the region of the predicted helix in the signal peptide near the helix in the mature PTH molecule (data not shown). Close approximation of these two helical regions may result in a change of conformation at the site of alternate cleavage that could lead to the use of the site unusually close to the hydrophobic core in the mutant precursor protein.

Comparative studies of various signals (Perlman and Halvorson, 1983; von Heinje, 1983, 1986a, 1986b) have attempted to define the structural limits of the signal sequence. However, the nature of the role played by sequences downstream of the signal peptide on the secretory process remains unclear. Studies with bacterial proteins have shown a variety of effects resulting from alterations of downstream sequences, including inefficient cleavage (Russell and Model, 1981), inaccurate cleavage (Dierstein and Wickner, 1986; Abrahmsen *et al.*, 1985), or no effect on processing (Haguenauer-Tsapis and Hinnen, 1984). Manipulation of sequences adjacent to signal sequences in eukaryotes had been infrequent. A major deletion in the beginning of the mature part of yeast acid phosphatase precursor resulted in minor impairment in processing *in vitro* (Takahara *et al.*, 1985). Recent reports have shown deletion of propeptides from apolipoprotein precursors can affect *in vitro* processing efficiency (Folz and Gordon, 1987) and choice of cleavage site (Folz and Gordon, 1986). The similarities of the abnormalities associated with the deleted forms of these precursors *in vitro* and prepro-PTH further support the suggestion that sequences immediately distal to the signal sequence importantly influence signal sequence function.

We suggest that one role of the propeptide is to act as an "adapter" between the signal sequence and the mature PTH molecule. Structure-function studies of PTH have demonstrated the vital role played by the serine at residue +1 of human PTH. Deletion of this serine residue or addition of another residue at position -1 dramatically decreases the bioactivity of PTH (Habener *et al.*, 1984). Since PTH acts by binding to cell surface receptors in several tissues Habener *et al.*, 1984), we can conclude that PTH and its receptor have evolved together in a way that places important constraints on the amino terminus of the mature PTH molecule. At the same time, the signal sequence of prepro-PTH had to evolve under the constraints dictated both by the requirements of signal peptidase and other roles played by the signal. Perhaps sequences needed at the amino terminus of PTH inevitably decrease the efficiency of translocation and accuracy of signal peptidase cleavage. The propeptide may therefore act in many proteins as an "adapter" when signal peptidase might otherwise cleave inefficiently, or might cleave in several sites or even solely at inappropriate sites. Implicit in such a formulation is a rejection of the alternative hypothesis that signal sequences can easily adapt to the constraints imposed by adjacent sequences. Experiments with further mutants of prepro-PTH and other secretory proteins will be required to better define these postulated constraints and the hypothesized limits of the adaptability of signal sequences.

In summary, we have directly tested the functional consequences of alteration of sequences immediately downstream from the signal on various aspects of signal function and on secretion from intact cells. We, of course, have not eliminated other potential roles for the pro-specific sequences. We have

² K. M. Wiren, L. Ivashkiv, P. Ma, M. W. Freeman, J. T. Potts, Jr., and H. M. Kronenberg, submitted for publication.

THIS PAGE BLANK (USPTO)

established, however, that deletion of the prosegment of prepro-PTH results in reduced efficiency of processing and in ambiguity in the sites selected by signal peptidase for cleavage *in vitro* and in intact cells. The prosegment therefore contributes to accurate and efficient signal function.

Acknowledgements—We are grateful to Drs. J. Hendy and J. O'Riordan for the PTH antisera provided. We thank Dr. J. Habener and G. Grynan for synthesizing the oligonucleotides. We are grateful to Dr. Y. Novotny for helpful discussions about secondary structure predictions. We also appreciate the expert secretarial support provided by Susan Rhodes and Brigid Pach.

REFERENCES

- Abrahmsen, L., Moks, T., Nilsson, B., Hellman, U., and Uhlen, M. (1985) *EMBO J.* **4**, 3901-3906
- Blobel, G., and Dobberstein, B. (1975a) *J. Cell Biol.* **67**, 835-851
- Blobel, G., and Dobberstein, B. (1975b) *J. Cell Biol.* **67**, 852-862
- Born, W., Freeman, M., Hendy, G. N., Rapoport, A., Rich, A., Potts, J. T., Jr., and Kronenberg, H. K. (1986) *Mol. Endocrinol.* **1**, 5-14
- Cepko, C. L., Roberts, B. E., and Mulligan, R. C. (1984) *Cell* **37**, 1053-1062
- Chou, P. Y., and Fasman, D. G. (1978) *Annu. Rev. Biochem.* **47**, 251-276
- Cohn, D. V., and Elting, J. (1983) *Recent Prog. Horm. Res.* **39**, 181-209
- Dierstein, R., and Wickner, W. (1986) *EMBO J.* **5**, 427-431
- Docherty, K., and Steiner, D. F. (1982) *Annu. Rev. Physiol.* **44**, 625-638
- Folz, R. J., and Gordon, J. I. (1986) *J. Biol. Chem.* **261**, 14752-14759
- Folz, R. J., and Gordon, J. I. (1987) *J. Biol. Chem.* **262**, 17221-17230
- Gilbert, W. (1978) *Nature* **271**, 501
- Graham, R., and Van der Eb, A. (1973) *Virology* **31**, 360-369
- Griffiths, G., and Simons, K. (1986) *Science* **234**, 438-443
- Guan, J. L., and Rose, J. K. (1984) *Cell* **37**, 779-787
- Guarente, L., Lauer, G., Roberts, T. M., and Ptashne, M. (1980) *Cell* **20**, 543-553
- Habener, J. F. (1979) *Endocrinology* (DeGroot, L. S., ed) Vol. 2, pp. 599-605, Grune and Stratton, New York
- Habener, J. F., Amherdt, M., Ravazzola, M., and Orci, L. (1979) *J. Cell Biol.* **80**, 715-731
- Habener, J. F., Rosenblatt, M., and Potts, J. T., Jr. (1984) *Physiol. Rev.* **64**, 985-1053
- Haguenauer-Tsapis, R., and Hinnen, A. (1984) *Mol. Cell Biol.* **4**, 2668-2675
- Heinrich, G., Kronenberg, H. M., Potts, J. T., Jr., and Habener, J. F. (1984) *J. Biol. Chem.* **259**, 3320-3329
- Hellerman, J. G., Cone, R. C., Potts, J. T., Jr., Rich, A., Mulligan, R. C., and Kronenberg, H. M. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 5340-5344
- Hendy, G. N., Kronenberg, H. M., Potts, J. T., Jr., and Rich, A. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 7365-7369
- Hebert, E., and Uhler, M. (1982) *Cell* **30**, 1-2
- Hollander, M., and Wolfe, D. A. (1973) *Nonparametric Statistical Methods*, p. 16, John Wiley & Sons, New York
- Hortin, G., and Boime, I. (1981) *J. Biol. Chem.* **256**, 1491-1494
- Hortsch, M., and Meyer, D. I. (1986) *Int. Rev. Cytol.* **102**, 215-242
- Jackson, R. C., and Blobel, G. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5598-5602
- Jorgensen, M. J., Cantor, A. B., Furie, B. C., Brown, C. L., Shoemaker, C. B., and Furie, B. (1987) *Cell* **48**, 185-191
- Kreil, G. (1981) *Annu. Rev. Biochem.* **50**, 317-48
- Kronenberg, H. M., Fennick, B. J., and Vasicek, T. J. (1983) *J. Cell Biol.* **96**, 1117-1119
- Laemmli, U. K. (1970) *Nature* **227**, 680-685
- Mann, R., Mulligan, R. C., and Baltimore, D. (1983) *Cell* **33**, 153-159
- Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560
- Mulligan, R. C., and Berg, P. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 2072-2076
- Noe, B. D., and Spiess, J. (1983) *J. Biol. Chem.* **258**, 1121-1128
- Orci, L., Ravazzola, M., Amherdt, M., Madsen, O., Perrelet, A., Vassalli, J.-D., and Anderson, R. G. W. (1986) *J. Cell Biol.* **103**, 2273-2281
- Parker, B. A., and Stark, G. R. (1979) *J. Virol.* **31**, 360-369
- Perlman, D., and Halvorson, H. O. (1983) *J. Mol. Biol.* **167**, 391-409
- Russell, M., and Model, P. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 1717-1721
- Shields, D., and Blobel, G. (1978) *J. Biol. Chem.* **253**, 3753-3756
- Steiner, D. F., and Clark, J. L. (1968) *Proc. Natl. Acad. Sci. U. S. A.* **60**, 622-629
- Steiner, D. F., Docherty, K., and Carrol, R. (1984) *J. Cell Biol.* **98**, 121-130
- Takahara, J., Hibler, D. W., Barr, P. J., Gerlt, J. A., and Inouye, M. (1985) *J. Biol. Chem.* **260**, 2670-2674
- Tashjian, A. H., Lomedico, M. E., and Maina, D. (1978) *Biochem. Biophys. Res. Commun.* **81**, 798-806
- Vasicek, T. J., McDevitt, B. E., Freeman, M. W., Fennick, B. J., Hendy, G. N., Potts, J. T., Jr., Rich, A., and Kronenberg, H. M. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 2127-2131
- von Heijne, G. (1983) *Eur. J. Biochem.* **133**, 17-21
- von Heijne, G. (1984) *J. Mol. Biol.* **173**, 243-251
- von Heijne, G. (1986a) *J. Mol. Biol.* **184**, 99-105
- von Heijne, G. (1986b) *Nucleic Acids Res.* **14**, 4683-4690
- Walter, P., Gilmore, R., and Blobel, G. (1984) *Cell* **38**, 5-8
- Weaver, C. A., Gordon, D. F., Kissel, M. S., Mead, D. A., and Kemper, B. (1984) *Gene (Amst.)* **28**, 319-329

THIS PAGE BLANK (USPTO)

A Deletion That Includes the Signal Peptidase Cleavage Site Impairs Processing, Glycosylation, and Secretion of Cell Surface Yeast Acid Phosphatase

ROSINE HAGUENAUER-TSAPIS[†]* AND ALBERT HINNEN

Friedrich Miescher Institut, Basel, Switzerland

Received 8 May 1984/Accepted 28 August 1984

We transformed *Saccharomyces cerevisiae* with a high-copy-number plasmid carrying either the wild-type gene coding for a repressible cell surface acid phosphatase or two modified genes whose products lack a 13- or 14-amino-acid segment spanning or immediately adjacent to the signal peptidase cleavage site. The wild-type gene product underwent proteolytic cleavage of the signal peptide, core glycosylation, and outer chain glycosylation. The deletion spanning the signal peptidase cleavage site led to an unprocessed protein. This modified protein exhibited core glycosylation, whereas its outer chain glycosylation was severely inhibited. Secretion of the deleted protein was impaired, and active enzyme accumulated within the cell. The deletion immediately adjacent to the signal peptidase cleavage site exhibited only a small decrease in the efficiency of processing and had no effect on the efficiency of secretion.

Studies on the biosynthesis of secretory proteins in prokaryotes and eucaryotes have revealed that most of the proteins are synthesized in vitro as precursors with 15 to 30 additional amino acids at the NH₂-terminal end of the molecule (31). This transient NH₂-terminal extension has been termed the signal peptide (6). It has been shown in eucaryotes that upon emergence of the signal peptide from the large ribosomal subunit, the ribosomal complex makes contact with the membrane of the endoplasmic reticulum (ER) by binding to the signal recognition particle (48), and the resulting complex binds in turn to the docking protein (28). The nascent protein is then cotranslationally translocated across the membrane of the ER, and the signal peptide is cleaved by signal peptidase (22).

In eucaryotes, protein secretion is often linked to protein glycosylation within the ER and the Golgi apparatus. The pathway of glycoprotein secretion in yeasts is similar to that described for higher eucaryotes (35), with only slight modifications. Isolation and characterization of numerous temperature-sensitive yeast secretory mutants blocked at various stages of this pathway (32, 33) led to the following model (15). The initial N-linked asparagine core glycosylation of secretory proteins probably occurs during translocation of the polypeptide through the membrane of the ER. Core oligosaccharides are then extended in a Golgi-like organelle by the addition of outer chain-sugar branches (3). Fully glycosylated proteins are then packed into vesicles which are transported to the bud, here they fuse with the plasma membrane and discharge their contents into the periplasmic space (32). This model should apply to repressible acid phosphatase. Although this enzyme has been studied extensively, it is not yet clear whether its final localization is in the periplasm (2) or in the cell wall (4, 25). Moreover, some of the enzyme is excreted (8).

Although genetic approaches have significantly contributed to our understanding of the essential features of signal

peptides in protein transport in *Escherichia coli* (17, 30), similar studies have only recently been reported for eucaryotic systems (16, 42). We have previously isolated and characterized two yeast acid phosphatase structural genes (*PHO5* and *PHO3*) arranged in tandem (29), both of which code for proteins secreted to the surface of the cell. The expression of the *PHO5* gene is repressed at high levels of phosphate, whereas the gene *PHO3* is constitutively expressed. Our initial sequencing data of the *PHO5* gene revealed a stretch of 20 neutral amino acids at the amino terminus of the protein, which strongly suggested the presence of a signal peptide. Amino acid sequence analysis of the mature protein (1, 46; unpublished data) was consistent with the existence of a signal peptide consisting of 17 amino acids. The presence of unique restriction sites flanking the presumed processing site (*BalI*, *SalI*, and two *KpnI* sites; see Fig. 1) enabled us to construct small deletions on either side of the processing site. We present here the results of an analysis of two deletions of *PHO5* [Δ *BalI-SalI*(Δ BS) and Δ *KpnI*(Δ K)] beginning, respectively, before and immediately after the signal peptidase cleavage site. We show that processing, glycosylation, and secretion of regulated acid phosphatase are affected in the Δ BS mutant, whereas processing and secretion are normal in the Δ K mutant. These results have been presented in part previously (19).

MATERIALS AND METHODS

Bacteria strains. *E. coli* HB101 (*hsdR hsdM leu pro recA*) was used for plasmid preparations. JM101 (Δ *lac pro supE thi F' traD36 proAB lac^rZAM15*) was used for M13 cloning.

Yeast strain. *Saccharomyces cerevisiae* GRF18 (α *his3-11 his3-15 leu2-3 leu2-112 can'*) was used.

Bacterial vectors. The phage M13mp8 was constructed by J. Messing (27).

Hybrid bacteria-yeast vectors. p29 is a pBR322 derivative; the DNA segment between *Bam*HI and *Pvu*II was replaced by a 3.9-kilobase (kb) *Bam*HI-*Hpa*I fragment containing *PHO5* and *PHO3* (W. Bajwa, B. Meyhack, H. Rudolph, A. M. Schweingruber, and A. Hinnen, submitted for publica-

* Corresponding author.

[†] Present address: Biochimie des Porphyrines, Université Paris VII, 75005 Paris, France.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Vortex mixer for 2 min at maximal speed as described by Ohashi et al. (34). Preparation of the samples for immunoprecipitation was then performed as described by the same authors, with slight modifications. Denatured cells and glass beads were centrifuged, and supernatant was removed. The sedimented proteins were solubilized by adding 1 ml of twofold-concentrated sample buffer (4% sodium dodecyl sulfate, 0.1 M Tris-hydrochloride (pH 6.8), 4 mM EDTA, 20% glycerol, 2% mercaptoethanol, 0.02% bromophenol blue). The mixture was neutralized with 1 M Tris base and heated for 4 min in a boiling bath. The suspension containing denatured solubilized proteins was then carefully removed from the glass beads. The glass beads were rinsed several times with Triton buffer. The rinses were combined with the first mixture and adjusted to a final volume of 20 ml (final sodium dodecyl sulfate concentration, 0.2%). For immunoprecipitation of the supernatant medium, 200- μ l samples of labeled cells were centrifuged (2 min; Eppendorf centrifuge), and the proteins of the supernatant medium were precipitated by the addition of 800 μ l of 20% trichloroacetic acid and 20 μ l of bovine serum albumin (20 mg/ml). The resulting pellet was resuspended in 0.5 ml of twofold-concentrated sample buffer, neutralized by 1 M Tris base, heated in a boiling bath for 4 min, and adjusted to a final volume of 10 ml of Triton buffer.

Immunoprecipitation. All Triton buffer samples resulting from cell-free translation or in vivo labeling were incubated for 30 min at room temperature with 100 μ l of a 10% suspension of heat-killed and glutaraldehyde-fixed *Staphylococcus aureus* cells and centrifuged for 20 min at 20,000 \times g. This step removed insoluble material and labeled proteins binding nonspecifically to bacterial cells. The pellet was discarded. Antiserum specific against purified acid phosphatase was added to the supernatants (10 μ l in the case of in vitro samples and supernatant medium samples; 20 μ l in the case of total cells), and overnight incubation was performed at room temperature. Fixed *S. aureus* cells (10-fold the volume of antiserum) were added, and the suspension was shaken for 1 h at room temperature. The *S. aureus* cells were then centrifuged, washed four times with 1 ml of Triton buffer, and extracted for 3 min at 95°C with 150 μ l of sample buffer. *S. aureus* cells were removed by centrifugation, and the supernatant was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis in 10% polyacrylamide. Gel electrophoresis was performed as described by Douglas and Butow (14), and the dried gels were analyzed by fluorography as described by Chamberlain (11).

RESULTS

Construction of high-copy-number plasmids carrying the PHO5 gene with a modified signal sequence. Small deletions in the DNA sequence coding for amino acids of the signal sequence, the adjacent mature PHO5 protein sequence, or both were used to analyze the involvement of this region in secretion and glycosylation. Figure 1 presents the nucleotide sequence at the 5' end of the PHO5 coding region (29) as well as the derived amino acid sequence. The amino acids at the NH₂ terminus show features characteristic of a signal sequence: there is a charged amino acid in position 3, followed by 14 neutral amino acids. The putative signal peptidase cleavage site is between amino acids Ala₁₇ and Gly₁₈ (1, 46). Two deletions have been constructed in plasmid p29 which carries the structural genes for PHO5 and PHO3 on a BamHI-HpaI fragment (see Fig. 2) with unique restriction sites (see Fig. 1). Deletion Δ K was obtained by digesting plasmid p29 with restriction endonuclease KpnI and religat-

ing the large DNA fragment. The new plasmid p29/PHO5/ Δ K is characterized by a deletion of 42 base pairs (14 amino acids). Deletion Δ BS was obtained by digesting plasmid p29 with restriction endonucleases BamHI and SalI, the 3'-recessed ends of the SalI site were filled in with Klenow DNA polymerase, and the resulting blunt ends of the large DNA fragment were ligated with T4 DNA ligase. This construction recreates a SalI restriction site. The new plasmid (p29/PHO5/ Δ BS) has a deletion of 39 base pairs (13 amino acids). The positions of the Δ K and Δ BS deletions are shown in Fig. 1. The Δ BS deletion removes three amino acids of the signal sequence, including the signal peptidase cleavage site and amino acids from the mature PHO5 amino acid sequence. The Δ K deletion leaves the signal sequence and the signal peptidase cleavage site intact but removes 14 amino acids from the mature PHO5 protein.

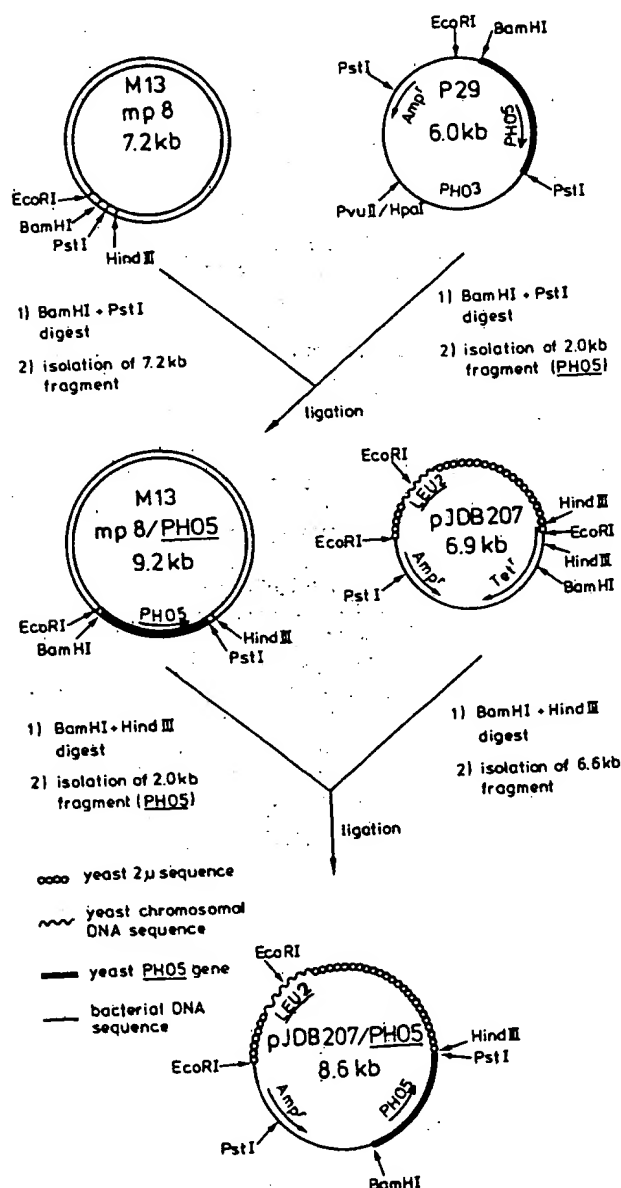


FIG. 2. Construction of expression plasmid pJDB207/PHO5 (BamHI-PstI).

THIS PAGE BLANK (USPTO)

TABLE 1. Acid phosphatase activity of intact yeast cells transformed by plasmids carrying wild-type or deleted *PHO5* gene

Yeast strain	Plasmid	Growth medium	Acid phosphatase activity (U/OD ₆₀₀) ^a	Derepression in low-P _i medium (fold)	Increase due to presence of plasmid in derepressed cells (fold)
GRF18	—	High P _i Low P _i	0.003 0.040		
GRF18	pJDB207/ <i>PHO5</i>	High P _i Low P _i	0.15 0.62	13	
GRF18	pJDB207/ <i>PHO5</i> ΔK	High P _i Low P _i	0.15 0.61	4	15
GRF18	pJDB207/ <i>PHO5</i> ΔBS	High P _i Low P _i	0.04 0.28	4 7	15 7

^a Yeast strain GRF18 (with or without the plasmids indicated in the table) was grown to stationary phase in minimal medium with glucose as the carbon source and was inoculated at an OD₆₀₀ of 0.05 into high-P_i or low-P_i medium. Growth was stopped when the OD₆₀₀ reached 1.5. Acid phosphatase activity of whole cells in their growth medium was then measured.

We introduced the wild-type *PHO5* gene into the high-copy-number yeast plasmid pJDB207 (5) which carries the *LEU2* gene. A 2.0-kb *Bam*HI-*Pst*I fragment which contains the whole *PHO5* gene (promoter and structural gene) was isolated from plasmid p29 and cloned into M13mp8 (27). This step results in the addition of a *Hind*III site close to the *Pst*I site at the end of the gene (Fig. 2). The plasmid pJDB207 was cleaved with *Bam*HI and *Hind*III and was then ligated with the 2.0-kb *Bam*HI-*Hind*III restriction fragment of M13mp8/*PHO5*. Plasmid pJDB207/*PHO5* was used in the construction of pJDB207/*PHO5* derivatives carrying the deletions ΔK and ΔBS. The 2.0-kb *Bam*HI-*Pst*I fragments of p29/*PHO5*ΔK and p29/*PHO5*ΔBS were isolated, and each was ligated to a mixture of the two 6.1- and 6.8-kb *Bam*HI-*Pst*I fragments resulting from complete *Bam*HI and partial *Pst*I digests of pJDB207/*PHO5*. Transformation of *E. coli* strain HB101 and selection for Amp^R clones enabled us to obtain plasmids with the expected deletions (pJDB207/*PHO5*ΔK and pJDB207/*PHO5*ΔBS). Restriction map analysis of the two deleted plasmids confirmed the extent of the deletions and the presence or absence of the expected restriction sites (data not shown).

Acid phosphatase activity in yeast cells transformed with plasmids carrying *PHO5* and the *PHO5*ΔK and *PHO5*ΔBS deletions. The three plasmids bearing the *PHO5* gene, either wild type or deleted (ΔK, ΔBS), were used to transform GRF18, a *Leu*⁻ wild-type yeast strain for the phosphatase genes. *Leu*⁺ clones were selected and tested for repressible acid phosphatase activity. None of the transformed strains exhibited any change in growth behavior. Acid phosphatase activity was measured in intact cells, since the substrate *p*-nitrophenylphosphate has free access to the cell wall and the periplasmic space. The activity was increased in the presence of any of the three plasmids, even after growth in high-P_i medium (Table 1). This suggests that the concentration of some regulatory component is not high enough to repress the system when the *PHO5* promoter is present on a high-copy-number plasmid. Under derepressed conditions in low-P_i medium, the presence of the plasmids pJDB207/*PHO5* and pJDB207/*PHO5*ΔK induced a 15-fold increase of acid phosphatase activity. Under these conditions, acid phosphatase was estimated to represent ca. 5% of the total protein as based on the specific activity of the purified enzyme (41).

Results also show that the ΔBS deletion led to a smaller

increase of activity as compared with the wild-type gene or the ΔK deletion (Table 1). To determine whether this smaller increase was caused by an incomplete secretion of the ΔBS deleted protein, the acid phosphatase activity was measured in toluene-permeabilized cells. Toluene treatment is known to impair plasma membrane permeability in *E. coli* (9) as well as in yeasts (43) and therefore leads to free access of the substrate to the cytoplasm. Transformants were shifted from minimal medium to low-P_i medium, and samples were taken at various times to permit the measurement of enzyme activity in intact whole cells, toluene-treated cells (Fig. 3A), and the supernatant medium (Fig. 3B). In the case of GRF18,

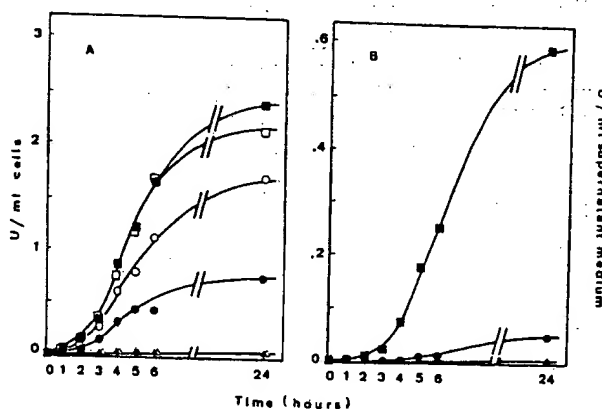


FIG. 3. Acid phosphatase activity in intact or toluene-treated cells (A) and in supernatant medium (B) after derepression of yeast cells transformed with *PHO5* and ΔBS plasmids. GRF18 nontransformed cells (▲, △) and cells transformed with *PHO5* (■, □) or ΔBS (●, ○) were grown up to stationary phase in minimal medium, harvested, and suspended at an OD₆₀₀ of 1 in low-P_i medium supplemented with histidine and leucine (20 μg/ml; untransformed cells) or histidine alone (transformed cells). Samples were withdrawn at various time points during derepression and tested for acid phosphatase activity of intact cells (A; closed symbols), or toluene-treated cells (A; open symbols). Samples were centrifuged (3,000 × g, 2 min), and acid phosphatase activity of the resulting supernatant was measured (B). At the end of the experiment, the OD₆₀₀ were 2.1, 2.3, and 2.2, respectively, for GRF18 control cells and cells transformed by *PHO5* and ΔBS plasmids.

THIS PAGE BLANK (USPTO)

the untransformed strain, and GRF18 transformed with the plasmid carrying the wild-type *PHO5* gene, the acid phosphatase activity of intact cells was identical to that of cells permeabilized with toluene (Fig. 3A). This was also the case for strains transformed by the plasmid pJDB207/*PHO5*/ΔK (data not shown). However, in the case of cells transformed by plasmid pJDB207/*PHO5*/ΔBS, the acid phosphatase activity of intact cells was half that measured in toluene-treated cells. Strains transformed by plasmid pJDB207/*PHO5*/ΔBS therefore produce an active phosphatase, only half of which reaches the cell envelopes even after 24 h of derepression. When GRF18 was transformed with the plasmid carrying the wild-type gene, ca. 25% of the acid phosphatase could be detected in the supernatant (Fig. 3A and B). This result was also obtained with the ΔK deleted acid phosphatase (data not shown). On the other hand, the ΔBS deletion led to a drastic decrease in the activity measured in the supernatant.

Processing of wild-type or deleted *PHO5* gene products in vivo and in vitro. The defect in secretion observed in the ΔBS deletion mutant was analyzed at the molecular level. Total RNA was prepared from the yeast strains transformed by plasmids carrying wild-type or deleted *PHO5* genes. In vitro translation products were analyzed after immunoprecipitation with an antibody reactive against acid phosphatase, followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and fluorography. The relative molecular mass observed for the in vitro *PHO5* gene product (Fig. 4, lane 1) was ca. 58 kilodaltons (kDa) which corresponds to the 60-kDa polypeptide observed by others (7, 40). Both estimates are slightly higher than those deduced from the DNA sequence (1; Bajwa et al., submitted for publication). The ΔK

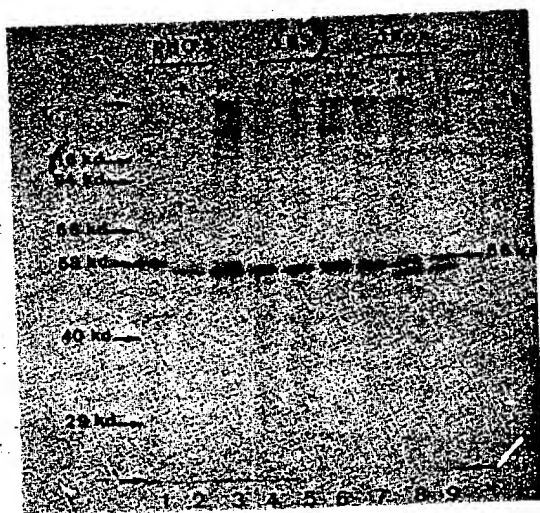


FIG. 4. In vivo and in vitro processing of *PHO5*, ΔK, and ΔBS gene products. Total RNA of transformed yeast cells grown and derepressed as described in the text was translated in a nuclease-treated reticulocyte lysate in the absence (–) or presence (+) of 40 μg of purified *E. coli* leader peptidase per ml. In vitro-translated proteins were submitted to immunoprecipitation, followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and fluorography. Yeast cells grown and derepressed in low- P_i , low- S_i medium were treated by tunicamycin and labeled in vivo for 30 min in the presence of 500 μCi of [35 S]methionine per ml. Labeled cells were submitted to trichloroacetic acid precipitation and immunoprecipitation as described in the text (lanes T). Arrows at top and bottom indicate origin and front of the gel, respectively. The mobilities of molecular weight standards are indicated by arrows on the left side of the figure.

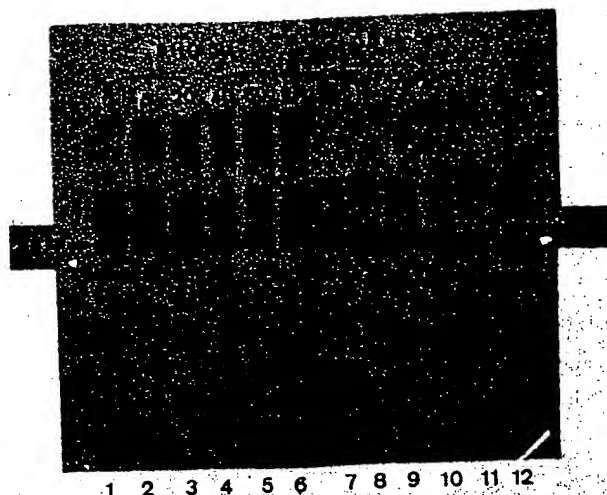


FIG. 5. Immunoprecipitation of acid phosphatase after pulse-chase labeling of intact yeast cells transformed with *PHO5* and ΔBS plasmids. Yeast cells grown and derepressed in low- P_i , low- S_i medium were submitted to a 5-min pulse labeling in the presence of 500 μCi of [35 S]methionine per ml, followed by a chase in the presence of 20 mM unlabeled methionine for the times indicated above each lane. Samples (200 μl) of total cells plus labeling medium were submitted to trichloroacetic acid precipitation and immunoprecipitation. Lanes 1 to 6, *PHO5*-transformed cells; lanes 7 to 12, ΔBS-transformed cells. Arrows are as defined in the legend to Fig. 4.

and the ΔBS in vitro products (Fig. 4, lanes 4 and 7) are both 56 kDa, i.e., ca. 2 kDa smaller than the wild-type in vitro product; this corresponds with the reduction predicted by the DNA sequence. The intensity of the signals obtained by immunoprecipitation reveals that the amount of translatable RNA specific for acid phosphatase is similar in all three transformed strains. In comparison, the signal obtained for the in vitro product of RNA prepared from the derepressed untransformed strain was ca. 20 times less intense (data not shown) and was not detectable under the experimental conditions described in the legend to Fig. 4.

In vivo proteolytic processing in all three transformed strains was studied in yeast strains labeled in the presence of tunicamycin, a compound which prevents *N*-asparagine-linked glycosylation of proteins (45) but not proteolytic processing of secretory proteins (13). The *PHO5* in vivo product was ca. 2 kDa smaller than the *PHO5* in vitro-translated product (Fig. 4, lane 3). Only trace amounts of the 58-kDa protein were detectable. Cells transformed with the *PHO5*/ΔK gene synthesized two polypeptides (Fig. 4, lane 9). One of these showed the same mobility as the corresponding in vitro product (56 kDa) and another was ca. 2 kDa smaller (54 kDa). In vivo and in vitro ΔBS products (Fig. 4, lanes 6 and 4) exhibited identical mobilities. These results may reflect complete, partial, and defective proteolytic processing, for the wild-type *PHO5* and ΔK and ΔBS polypeptides, respectively, which were formed in vivo in the presence of tunicamycin. A confirmation of these results was obtained by in vitro processing with purified *E. coli* leader peptidase (49). This enzyme removes the signal peptides from a wide variety of preproteins in bacteria and some preproteins in eucaryotes. The presence of leader peptidase during in vitro translation led to total, partial, or no proteolytic processing, for wild-type *PHO5*, ΔK, and ΔBS in vitro products, respectively (Fig. 4; lanes 2, 8, and 5).

THIS PAGE BLANK (USPTO)

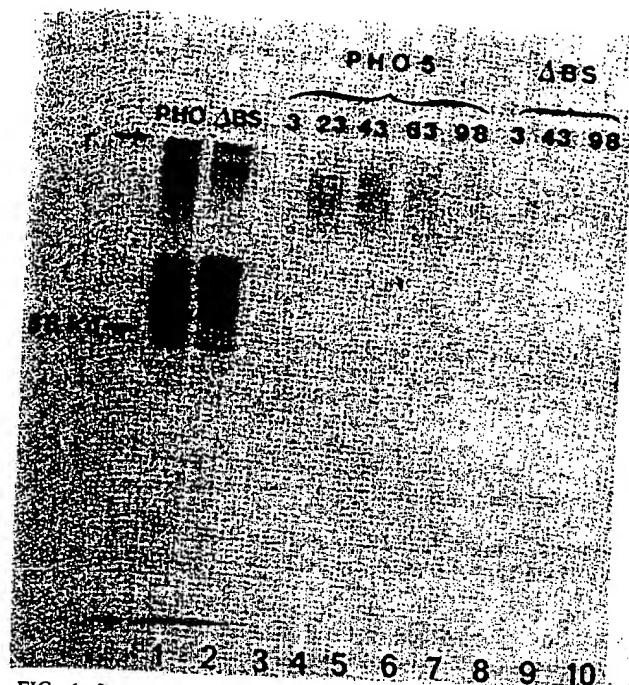


FIG. 6. Immunoprecipitation of acid phosphatase in the supernatant medium after a pulse-chase labeling experiment of yeast cells transformed with *PHO5* or Δ BS plasmid. Cells were labeled for 5 min and then chased with unlabeled methionine as described in the legend to Fig. 5. At the time points (min) indicated above the lanes, 200- μ l samples were centrifuged ($12,000 \times g$, 2 min), and the resulting supernatants were submitted to trichloroacetic acid precipitation and immunoprecipitation as described in the text. Lanes at left marked PHO and Δ BS correspond to 100 μ l of total cells plus supernatant medium after 90 min of labeling. Arrows are as defined in the legend to Fig. 4.

In vivo glycosylation of overproduced wild-type *PHO5* and *PHO5*/ Δ BS gene products. The effect of the Δ BS mutation on the glycosylation pattern of acid phosphatase was studied with *in vivo* labeling. Cells transformed by wild-type *PHO5* and *PHO5*/ Δ BS plasmids were derepressed for acid phosphatase and submitted to a 5-min pulse in the presence of [35 S]methionine, followed by a long chase with unlabeled methionine (up to 90 min). Immunoprecipitation was performed in parallel on whole cells (Fig. 5) and on the supernatant (Fig. 6, lanes 3 to 10). The same culture was also labeled for a period of 90 min (Fig. 6, lanes 1 and 2). The 5-min labeling of *PHO5*-transformed cells led to an accumulation of unprocessed and unglycosylated acid phosphatase (58 kDa; Fig. 5, lane 1); processed and unglycosylated enzyme (56 kDa) was hardly detectable. At the beginning of the chase, several discrete bands of increasing molecular weight (up to 80 kDa) were clearly observed, this suggests various stages in core glycosylation. As the chase progressed, the quantity of both unprocessed and core-glycosylated immunoreactive polypeptides decreased, whereas high-molecular-weight species (100 to 140 kDa) showed a concomitant increase. These high-molecular-weight species most probably correspond to outer chain glycosylated polypeptides (15). Although very low amounts of the 56-kDa protein and of degraded proteins were detected by immunoprecipitation (Fig. 6, lane 1), the kinetics of the overall process suggest that the uncleaved unglycosylated polypeptide might be the precursor of the glycosylated polypeptides. Only the 100- to 140-kDa species were excreted into the medium in which they appeared after 23 min of chase (Fig. 6, lanes 3 to 7).

Glycosylation and secretion of Δ BS deleted acid phosphatase were significantly different from the corresponding events with the wild-type protein. A 5-min labeling of cells only led to the formation of unprocessed and unglycosylated polypeptide (Fig. 5, lane 7). A subsequent chase led to a progressive decrease in intensity of the unglycosylated polypeptide band, with an increase in core glycosylated polypeptides of mainly 80 kDa (Fig. 5, lanes 8 to 12). The amount of proteolytically degraded material was much less than in the *PHO5* case. For the *PHO5*/ Δ BS-transformed cells, therefore, the 56-kDa polypeptide appeared as the precursor of the 80-kDa species. Fully glycosylated polypeptides were barely detectable in the pulse-chase experiment (Fig. 5, lanes 7 to 12) but were clearly visible after a 90-min labeling (Fig. 6, lane 2). They were characterized by a narrower range of molecular weights (120 kDa) than the corresponding wild-type glycoproteins. It might be argued that the high-molecular-weight glycosylated bands represent a wild-type acid phosphatase coded for by the chromosomal copy of the *PHO5* gene. However, this would be inconsistent with the relative abundance of these bands after a long labeling period, the apparently narrow range of their molecular weights, the absence of excretion of the corresponding proteins (Fig. 6, lanes 8 to 10), and the total absence of the 58-kDa *in vitro* product after *in vitro* translation of RNA from *PHO5*/ Δ BS-transformed strains (Fig. 4, lane 4).

DISCUSSION

We have transformed the yeast organism *S. cerevisiae* with a high-copy-number plasmid carrying a repressible acid phosphatase gene (*PHO5*) either as a wild-type allele or as one of two mutant forms with small deletions in or near the signal peptide. The two mutations exhibit specific effects on processing, glycosylation, and secretion.

We have shown that the wild-type *PHO5* protein is produced *in vitro* as a precursor which is proteolytically processed to produce the mature form as has been shown for yeast invertase (10, 37). *In vivo* labeling of yeast strains carrying multiple copies of the wild-type *PHO5* gene led to the transient formation of a 58-kDa polypeptide which presumably represents the unprocessed and unglycosylated phosphatase precursor polypeptide. This 58-kDa preprotein is also detectable after inhibition of glycosylation by tunicamycin. This is unusual since in the case of secreted eucaryotic proteins, cotranslational proteolytic processing is believed to precede chain completion (21). The presence of acid phosphatase preprotein is probably a consequence of an overproduction of the protein, leading to the saturation of some component of the processing machinery such as the signal peptidase.

In vitro processing by bacterial signal peptidase under appropriate experimental conditions was found to be effective for wild-type *PHO5*-coded polypeptide, ineffective for *PHO5*/ Δ BS-coded polypeptide, and partially effective for *PHO5*/ Δ K-coded polypeptide. This bacterial enzyme has already been shown to be capable of cleaving some eucaryotic preproteins (49). Recently, it has also been shown that an antibody specific against bacterial leader peptidase specifically recognizes a yeast microsomal protein (39). The DNA sequence of the wild-type gene and the amino acid sequence of the mature protein (1, 46) suggest that the cleavage site lies between amino acids 17 and 18. This is consistent with the inability of leader peptidase to cleave the *PHO5*/ Δ BS-coded polypeptide, which still possesses the first 14 of the 17 signal peptide amino acids, and is also

THIS PAGE BLANK (USPTO)

consistent with its ability to cleave the *PHO5*/ Δ K-coded protein whose deletion begins at two amino acids after the putative cleavage site. Although these results do not specify exactly the features which characterize a signal peptidase cleavage site, they do suggest that one essential feature for specific cleavage resides in the five amino acids which separate the beginning of the two deletions. Of these, the alanine at the carboxy terminus of the signal peptide may play a critical role: it is in this position in numerous other signal peptides (30). The reduced efficiency of cleavage of the signal peptide displayed by the Δ K deleted protein suggests that the beginning of the mature protein sequence participates in the interaction between nascent acid phosphatase and signal peptidase as has been shown for an M13 coat protein mutant (30).

Within the first 5 min of *in vivo* synthesis, Δ BS deleted protein exhibits an increase of molecular weight from 56 kDa up to 80 kDa. As in the case of wild-type protein, this increase is completely prevented by tunicamycin, an antibiotic that blocks *N*-asparagine-linked glycosylation by inhibiting the formation of dolichyl-*N*-acetyl-glucosaminylpyrophosphate. Therefore, Δ BS deleted protein, which still contains two-thirds of the hydrophobic portion of the signal peptide, undergoes core glycosylation. This probably reflects the ability of the mutated protein to cross the ER membrane, since *N*-asparagine-linked core glycosylation is thought to occur on the cisternal surface of the ER (20, 44). This situation is very different from that observed with an influenza virus hemagglutinin lacking either 11 amino acids from the 16 amino acids of its signal peptide (42) or the whole signal peptide (16). In these cases, an intracellular accumulation of unglycosylated forms of these deleted proteins has been demonstrated.

In the Δ BS deletion, the absence of cleavage of the signal peptide severely impairs outer chain glycosylation and leads to an incomplete secretion of the protein: 50% of the mutated protein is found as an active enzyme inside the cell, even after 24 h of derepression; the other 50%, however, is secreted to the cell surface, although in a form which cannot cross the cell wall. This result is compatible with the observation that, in *E. coli*, half of a mutated prolipoprotein in which the signal peptide is not cleaved still reaches its destination in the bacterial outer membrane (23, 24). In this case, the absence of removal of the signal peptide leads to a decrease in the rate of secretion. Establishing whether this is also the case for the Δ BS deleted acid phosphatase should further our understanding of glycoprotein export as should the precise localization and characterization of the extracellular and intracellular forms of the Δ BS deleted proteins.

ACKNOWLEDGMENTS

We thank W. Wickner for kindly providing us with purified bacterial leader peptidase, G. Schatz for kindly providing us with acid phosphatase antibody, H. Rudolph, B. Meyhack, W. Bajwa, H. Riezman, M. Suissa, R. Hay, R. Schekman, and G. Schatz for numerous helpful discussions, R. Labbe and V. Norris for critical reading of the manuscript, and D. Martin for excellent technical assistance.

R.H.-T. was the recipient of a European Molecular Biology Organization long-term postdoctoral fellowship.

LITERATURE CITED

1. Arima, K., T. Oshima, I. Kubota, N. Nakamura, T. Mizunaga, and A. Toh-e. 1983. The nucleotide sequence of the yeast *PHO5* gene: a putative precursor of repressible acid phosphatase. *Nucleic Acids Res.* 11:1657-1672.
2. Arnold, W. N. 1972. Location of acid phosphatase and β -fructofuranosidase within yeast cell envelopes. *J. Bact.* 112:1346-1352.
3. Ballou, C. E. 1976. Structure and biosynthesis of the mannan component of the yeast cell envelope. *Adv. Microb. Physiol.* 14:93-158.
4. Bauer, H., and E. Sigarlakie. 1973. Cytochemistry on ultrathin frozen sections of yeast cells. Localization of acid and alkaline phosphatase. *J. Microsc.* 99:205-218.
5. Beggs, J. D. 1981. Gene cloning in yeast, p. 175-203. In R. Williamson (ed.), *Genetic engineering*, vol. 2. Academic Press, London.
6. Blobel, G., and B. Dobberstein. 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and non processed nascent immunoglobulin chains on membrane bound ribosomes of murine melanoma. *J. Cell Biol.* 67:835-851.
7. Bostian, K. A., J. M. Lemire, L. E. Cannon, and H. O. Halvorson. 1980. *In vitro* synthesis of repressible yeast acid phosphatase: identification of multiple mRNAs and products. *Proc. Natl. Acad. Sci. U.S.A.* 77:4504-4508.
8. Bostian, K. A., J. M. Lemire, and H. O. Halvorson. 1983. Physiological control of repressible acid phosphatase gene transcripts in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 3:839-853.
9. Buttin, G. 1963. Mecanismes regulateurs dans la biosynthèse des enzymes du métabolisme du galactose chez *Escherichia coli* K12. I. La biosynthèse induite de la galactokinase et l'induction simultanée de la séquence enzymatique. *J. Mol. Biol.* 7:164-182.
10. Carlson, M., R. Taussig, S. Kustu, and D. Bostein. 1983. The secreted form of invertase in *Saccharomyces cerevisiae* is synthesized from mRNA encoding a signal sequence. *Mol. Cell. Biol.* 3:439-447.
11. Chamberlain, J. P. 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water soluble fluor: sodium salicylate. *Anal. Biochem.* 98:132-135.
12. Clewell, D. B. 1972. Nature of Col E₁ plasmid replication in *Escherichia coli* in the presence of chloramphenicol. *J. Bacteriol.* 110:667-676.
13. Dobberstein, B., H. Garoff, and G. Warren. 1979. Cell free synthesis and membrane insertion of mouse H-2D^a histocompatibility antigen and β_2 microglobulin. *Cell* 17:759-769.
14. Douglas, M., and R. A. Butow. 1976. Variant forms of mitochondrial translation products in yeast: evidence for location of determinants on mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* 73:1083-1086.
15. Esmon, B., P. Novick, and R. Schekman. 1981. Compartmentalized assembly of oligosaccharides on exported glycoproteins in yeast. *Cell* 25:451-460.
16. Gething, M. J., and J. Sambrook. 1982. Construction of influenza haemagglutinin genes that code for intracellular and secreted forms of the protein. *Nature (London)* 300:598-602.
17. Hall, M. N., and T. J. Silhavy. 1981. Genetic analysis of the major outer membrane proteins of *Escherichia coli*. *Annu. Rev. Genet.* 15:91-142.
18. Hinnen, A., J. B. Hicks, and G. R. Fink. 1978. Transformation of yeast. *Proc. Natl. Acad. Sci. U.S.A.* 75:1929-1933.
19. Hinnen, A., B. Meyhack, and R. Tsapis. 1983. High expression and secretion of foreign proteins in yeast, p. 157-163. In M. Korhola and E. Väisänen (ed.), *Gene expression in yeast. Proceeding of the Alko yeast Symposium*, Helsinki, vol. 1. Foundation for Biotechnical and Industrial Fermentation Research, Helsinki.
20. Katz, F. N., J. E. Rothman, V. R. Lingappa, G. Blobel, and H. F. Lodish. 1977. Membrane assembly *in vitro*: synthesis, glycosylation, and asymmetric insertion of a transmembrane protein. *Proc. Natl. Acad. Sci. U.S.A.* 74:3278-3282.
21. Kreil, G. 1981. Transfer of proteins across membranes. *Annu. Rev. Biochem.* 50:317-348.
22. Kreil, G., C. Mollay, R. Kashnitz, U. Vilas, and L. Haiml. 1980. Prepromellitin specific cleavage of the prepeptide and the propeptide *in vitro*. *Ann. N.Y. Acad. Sci.* 343:338-346.

THIS PAGE BLANK (USPTO)

23. Lin, J. J. C., H. Kanazawa, J. Ozols, and H. C. Wu. 1978. An *Escherichia coli* mutant with an amino acid alteration within the signal sequence of outer membrane prolipoprotein. *Proc. Natl. Acad. Sci. U.S.A.* 75:4891-4895.
24. Lin, J. J. C., H. Kanazawa, and H. C. Wu. 1980. Assembly of outer membrane lipoprotein in an *Escherichia coli* mutant with a single amino acid replacement within the signal sequence of prolipoprotein. *J. Bacteriol.* 141:550-557.
25. Linnemans, W. A. M., P. Boer, and P. F. Elbers. 1977. Localization of acid phosphatase in *Saccharomyces cerevisiae*: a clue to cell wall formation. *J. Bacteriol.* 131:638-644.
26. Mandel, M., and A. Higa. 1970. Calcium dependent bacteriophage DNA infection. *J. Mol. Biol.* 53:159-162.
27. Messing, J. 1981. M₁₃ mp2 and derivatives: a molecular cloning system for DNA sequencing, strand specific hybridization, and in vitro mutagenesis, p. 143-153. In A. G. Walton (ed.), *Recombinant DNA. The Third Cleveland Symposium on Macromolecules*. Elsevier, Amsterdam.
28. Meyer, D. I., E. Krause, and B. Dobberstein. 1982. Secretory protein translocation across membranes: the role of the "docking protein." *Nature (London)* 297:647-650.
29. Meyhack, B., W. Bajwa, H. Rudolph, and A. Hinnen. 1982. Two yeast acid phosphatase structural genes are the result of a tandem duplication and show different degrees of homology in their promoter and coding sequence. *EMBO J.* 1:675-680.
30. Michaelis, S., and J. Beckwith. 1982. Mechanism of incorporation of cell envelope proteins in *Escherichia coli*. *Annu. Rev. Microbiol.* 36:435-455.
31. Milstein, C., G. G. Brownlee, T. M. Harrison, and M. B. Matthews. 1972. A possible precursor of immunoglobulin light chains. *Nature (London) New Biol.* 239:117-120.
32. Novick, P., S. Ferro, and R. Schekman. 1981. Order of events in the yeast secretory pathway. *Cell* 25:461-469.
33. Novick, P., C. Fields, and R. Schekman. 1980. Identification of 23 complementation groups required for posttranslational events in the yeast secretory pathway. *Cell* 21:205-215.
34. Ohashi, A., J. Gibson, I. Gregor, and G. Schatz. 1982. Import of proteins into mitochondria. The precursor of cytochrome c₁ is processed in two steps, one of them heme dependent. *J. Biol. Chem.* 257:13042-13047.
35. Palade, G. 1975. Intracellular aspects of the process of protein synthesis. *Science* 189:347-357.
36. Pelham, H. R. B., and R. J. Jackson. 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur. J. Biochem.* 67:247-256.
37. Perlman, D., and H. O. Halvorson. 1981. Distinct repressible mRNAs for cytoplasmic and secreted yeast invertase are encoded by a single gene. *Cell* 25:525-536.
38. Reid, G. A., T. Yonetani, and G. Schatz. 1982. Import of proteins into mitochondria. Import and maturation of the mitochondrial intermembrane space enzymes cytochrome b₂ and cytochrome c peroxidase in intact yeast cells. *J. Biol. Chem.* 257:13068-13074.
39. Riezman, H., T. Hase, A. P. G. M. Van Loon, L. A. Grivell, K. Suda, and G. Schatz. 1983. Import of proteins into mitochondria: a 70 kilodalton outer membrane protein with a large carboxy-terminal deletion is still transported to the outer membrane. *EMBO J.* 2:2161-2168.
40. Rogers, D. T., J. M. Lemire, and K. A. Bostian. 1982. Acid phosphatase polypeptides in *Saccharomyces cerevisiae* are encoded by a differentially regulated multigene family. *Proc. Natl. Acad. Sci. U.S.A.* 79:2157-2161.
41. Schweingruber, A. M., and E. Schweingruber. 1982. Differential regulation of the active and inactive forms of *Saccharomyces cerevisiae* acid phosphatase. *Mol. Gen. Genet.* 187:107-111.
42. Sekikawa, K., and C. J. Lai. 1983. Defects in functional expression of an influenza virus haemagglutinin lacking the signal peptide sequences. *Proc. Natl. Acad. Sci. U.S.A.* 80:3563-3567.
43. Serrano, R., J. M. Gancedo, and C. Gancedo. 1973. Assay of yeast enzymes in situ: a potential tool in regulation studies. *Eur. J. Biochem.* 34:479-482.
44. Snider, M. D., and P. W. Robbins. 1982. Transmembrane organization of protein glycosylation: mature oligo-saccharide-lipid is located on the luminal side of microsomes from chinese hamster ovary cells. *J. Biol. Chem.* 257:6796-6801.
45. Takatsuki, A., K. Khono, and G. Tamura. 1975. Inhibition of biosynthesis of polyisoprenol sugars in chick embryo microsomes by tunicamycin. *Agric. Biol. Chem.* 39:2089-2091.
46. Thill, G. P., R. A. Kramer, K. J. Turner, and K. A. Bostian. 1983. Comparative analysis of the 5'-end regions of two repressible acid phosphatase genes in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 3:570-579.
47. Toh-e, A., Y. Ueda, S.-I. Kakimoto, and Y. Oshima. 1973. Isolation and characterization of acid phosphatase mutants in *Saccharomyces cerevisiae*. *J. Bacteriol.* 113:727-738.
48. Walter, P., and G. Blobel. 1981. Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of *in vitro* assembled polysomes synthesizing secretory proteins. *J. Cell Biol.* 91:551-556.
49. Wolfe, P. B., P. Silver, and W. Wickner. 1982. The isolation of homogenous leader peptidase from a strain of *Escherichia coli* which overproduces the enzyme. *J. Biol. Chem.* 257:7898-7902.

THIS PAGE BLANK (USPTO)

Roles of the signal peptide and mature domains in the secretion and maturation of the neutral metalloprotease from *Streptomyces cacaoi*

Su-Chih CHANG, Mei-Han SU and Yan-Hwa Wu LEE*

Institute of Biochemistry, National Yang-Ming University, Taipei 112, Taiwan, Republic of China

The neutral metalloprotease (Npr) of *Streptomyces cacaoi* is synthesized as a prepro-Npr precursor form consisting of a secretory signal peptide, a propeptide and the mature metalloprotease. The maturation of Npr occurs extracellularly via an autoproteolytic processing of the secreted pro-Npr. The integrity of the propeptide is essential for the formation of mature active Npr but not for its secretion [Chang, Chang and Lee (1994) *J. Biol. Chem.* 269, 3548–3554]. In this study we investigated whether the secretion and maturation of Npr require the integrity of its signal peptide region and mature protease domain. Five signal peptide mutants were generated, including the substitution mutations at the positively charged region (mutant IR6LE), the central hydrophobic region (mutants GI19EL and G19N), the boundary of the hydrophobic core-cleavage region (mutant P30L) and at the residues adjacent to the signal peptidase cleavage site (mutant YA33SM). All these lesions delayed the export of Npr to the growth medium and also resulted in a 2–10-fold decrease in Npr export. The most severe effect was noted in mutants GI19EL and P30L. When these signal peptide mutations

were fused separately with the propeptide lacking the Npr mature domain, the secretory defect on the propeptide was also observed, and this impairment was again more severely expressed in mutants GI19EL and P30L. Thus the Npr signal peptide seems to have more constraints on the hydrophobic core region and at the proline residue within the boundary of the hydrophobic core-cleavage site. Deletion mutations within the C-terminal mature protease domain that left its active site intact still blocked the proteolytic processing of mutant precursor forms of pro-Npr, although their secretions were unaffected. These results, together with our previous findings, strongly suggest that the signal peptide of Npr plays a pivotal role in the secretion of both Npr and the propeptide, but not in the maturation of Npr. On the contrary, the integrity of mature domain and propeptide is not critical for secretion of the Npr derivative but is essential for the formation of a functional Npr. Therefore the secretion and maturation of Npr are dependent on the integrity of the signal peptide, propeptide and mature protease domains, and the roles of these domains in this regard are functionally distinct.

INTRODUCTION

Virtually all prokaryotic secretory proteins destined for extracellular or membrane compartments are characterized by the presence of a signal peptide or leader sequence located at the N-terminus, although some internal signal peptides exist [1,2]. Analyses of a large number of the signal peptides reveal certain common structural features [2]: (1) all possess from one to three basic residues at the N-terminus followed by a hydrophobic segment of 10–18 residues; (2) residues capable of forming β -turns (e.g. proline and glycine) are often near the C-terminal end of this hydrophobic sequence; (3) specific residues with small side-chains (e.g. serine, alanine and glycine) immediately preceding the cleavage site constitute a recognition site for the signal peptidase. The relative importance of these structural features in secretion and their interactions with the components of the secretory apparatus of the cell have been the subject of extensive studies, largely in *Escherichia coli* (reviewed in [1–4]). In comparison, however, the secretory mechanisms of Gram-positive bacteria such as *Bacillus* or *Streptomyces* have received insufficient characterization. Few signal peptide sequence mutants in these two genera have been reported [5–8].

Our laboratory has been interested in the secretory system of *Streptomyces*. We have chosen MelC1 in the melanin gene (*melC*) operon and neutral metalloprotease (Npr) as model

systems for the study of *Streptomyces* secretory proteins. MelC1 has an N-terminal signal peptide sequence and acts as the molecular chaperone for apotyrosinase (MelC2); it has a dual role in secretion and *trans*-activation of apo-tyrosinase [8–11]. Npr is a neutral metalloprotease of *Streptomyces cacaoi* [12] and is synthesized as a precursor protein (prepro-Npr) consisting of an N-terminal signal peptide (34 residues) followed by a propeptide (171 residues) and a mature protease domain (345 residues) [12,13]. This protease has zinc-binding ligands located at His-202, His-206 and Glu-240 residues of the mature protease domain that are indispensable to the protease activity and proteolytic processing of the pro-Npr derivative [13]. On the basis of our previous studies [13,14] we propose a model (Figure 1) for the extracellular secretion and maturation of Npr. Our assumption is that the signal peptide guides the prepro-Npr precursor to pass through the membrane via the regular co- or post-translational transport route (step 1). During translocation the signal peptide of the prepro-Npr is removed by signal peptidase, resulting in the release of pro-Npr into the medium (step 2). Soon after arrival in the medium the propeptide promotes the correct folding of the Npr mature domain to acquire the active conformation, which then subsequently catalyses the autoproteolytic cleavage of pro-Npr propeptide (step 3). The released propeptide is further degraded, to form the mature protease (step 4). This maturation mechanism is sup-

Abbreviations used: *melC*, melanin gene locus; Npr, neutral metalloprotease; *npr*, gene encoding Npr; prepro-Npr (p60), preproenzyme of Npr; pro-Npr (p57), proenzyme of Npr; Pro, propeptide of Npr; TSB, tryptic soy broth.

* To whom correspondence should be addressed.

THIS PAGE BLANK (USPTO)

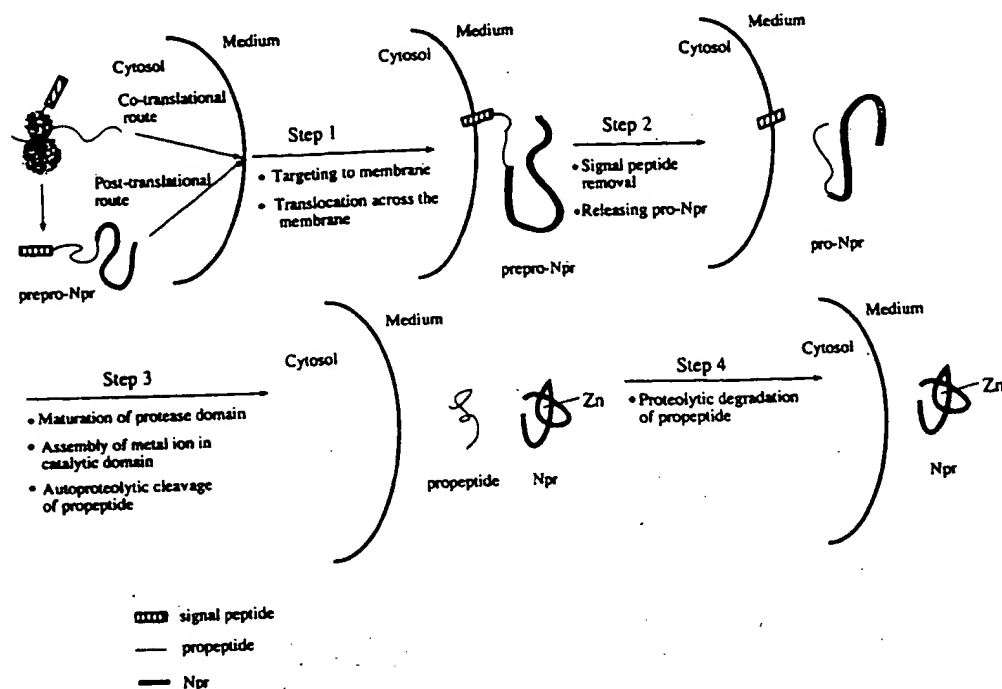


Figure 1 Proposed model for maturation and secretion of Npr

The signal peptide directs prepro-Npr across the membrane via the regular co- or post-translational transport route (step 1), where the signal peptide is removed, resulting in the release of pro-Npr (step 2). In the medium, folding of the mature domain of the protease occurs, including the metal ion (Zn)-binding site guided by the propeptide (step 3). A propeptide and mature protease are released from the precursor by autoproteolytic cleavage between the propeptide and the protease region. The propeptide is further degraded, leaving the mature Npr (step 4).

ported by the detection of pro-Npr in the medium of cells harbouring a mutant *npr* gene defective in the zinc-binding site [13]. However, in the cells harbouring a wild-type *npr* gene only the mature Npr can be detected in the medium [12–14], suggesting that the processes of propeptide-mediated folding and auto-processing of pro-Npr occur spontaneously.

An important feature of this model (Figure 1) is the involvement of the propeptide region in the maturation and secretory pathway of Npr. Earlier work conducted in our laboratory on the role of propeptide in Npr has shown that its complete removal inhibits the production of functional Npr [14]. Partial deletion of the propeptide region while leaving an intact mature Npr sequence resulted in the secretion of inactive, apparently unstable, truncated forms of pro-Npr [14]. Thus the integrity of the propeptide is essential for the maturation and correct processing of Npr but not for the secretion of Npr or its derivatives. This feature distinguishes the mechanism of maturation and secretion of Npr from those of subtilisin and the thermolysin-like neutral protease of *Bacillus*. In subtilisin the folding and the secretion functions of the propeptide sequences are not separable, because only properly folded and processed forms of subtilisin are secreted efficiently [15,16]. In the neutral protease of *Bacillus cereus*, the intact propeptide region is not necessary for the processing of *B. cereus* neutral protease precursor [17]. Therefore the protein export and maturation of these three proteases of Gram-positive bacteria seem to be considerably different with regard to the role of the propeptide.

Apart from the role of the propeptide, the relative importance of the mature domain and the signal peptide in Npr maturation and secretion remains unknown. It is also unclear to what extent transport of the precursor of Npr shares mechanistic features with the transport of the precursor protein to the periplasm of

Gram-negative bacteria, which in general is signal-peptide-dependent and requires that the nascent protein be in an unfolded state [18,19]. It is also important to know whether the signal peptide sequences in these two streptomycete proteins, MelC1 and Npr, have the same features in regard to their export function. Previously we reported the effects of mutations in the MelC1 signal peptide on its secretion and *trans*-activation functions [8].

In this study we examined the effects on the maturation and secretion of Npr by various mutations within the signal peptide sequence and mature region. Our results indicate that the integrity of the Npr signal peptide is crucial for the secretion of Npr and the propeptide. However, the integrity of the mature domain is not required for secretion of the Npr derivative, but is essential for the maturation of a functional Npr.

EXPERIMENTAL

Bacterial strains, plasmids, culturing conditions and recombinant DNA techniques

Streptomyces lividans TK64 (*pro*-2, *str*-6) [20] was the host for harbouring and expressing the cloned *npr* gene. The *Streptomyces* plasmid pSP105 harbouring the cloned *npr* gene has been described previously [13]. Plasmids pMUL1 and pMUS1 were constructed by subcloning the 5.4 and 0.93 kb *Pst*I–*Sac*I fragments of pSP105 respectively into the corresponding sites on vector pUC19 [14]. pSelect-BS carrying the 0.38 kb *Bam*HI–*Sac*I fragment containing the signal peptide and propeptide regions of *npr* gene was constructed by inserting this fragment into the corresponding sites on pSelect®-1 (Promega). Plasmids pVU1200 and pVU5100 were generated by inserting the 1.2 or 5.1 kb *Pvu*II fragments of pSP105 into the *Pvu*II site of the vector pUC19

THIS PAGE BLANK (USPTO)

	Charged region	Hydrophobic region	Cleavage site
wild type	MPMFRIRLPK PAALIAAGGIGACIATVAVPSAYAAAP*	↑
IR6LE	MPMFRLELP KPAALIAAGGIGACIATVAVPSAYAAAP		
GI19EL	MPMFRIRLP KPAALIAAGGIGACIATVAVPSAYAAAP		
G19N	MPMFRIRLP KPAALIAAGGIGACIATVAVPSAYAAAP		
P30L	MPMFRIRLP KPAALIAAGGIGACIATVAVPSAYAAAP		
YA33SM	MPMFRIRLP KPAALIAAGGIGACIATVAVPSAYAAAP		

Figure 2 Sequence of the wild-type and mutant signal peptides of Npr

The amino acid sequence of the wild-type Npr signal peptide is shown as standard single-letter abbreviations. The N-terminal positively charged region is boxed. The asterisks indicate the hydrophobic core region. The arrow marks the cleavage site by signal peptidase [13]. The residues substituted in each mutant are underlined.

respectively [13]. Culturing and gene manipulations of *E. coli* and *Streptomyces* were by the methods of Sambrook et al. [21] and Hopwood et al. [22].

Site-specific mutagenesis of the signal peptide region of *npr*

The oligonucleotide-directed site-specific mutagenesis of the *npr* signal peptide sequence was performed on plasmid pSelect-BS, using the Altered Site *In Vitro* Mutagenesis System (Promega Co.) in accordance with the specifications of the manufacturer. Oligonucleotides (designated by the position of the mutated amino acid and the amino acid before and after the mutation) are listed below. Mutated bases are shown in bold, and new restriction sites created after mutation are also indicated.

IR6LE	5'-TTTCGGCA ACTCGAGT CGGAACAT-3'
	XhoI
GI19EL	5'-ATGCAGGCGCCGAGCTCGCCCGCCGC-3'
G19N	5'-AGGCGCCTATGTTGCCCGCCGC-3'
P30L	5'-AGGCGGACAGTACGGCG-3'
YA33SM	5'-CGGGGGCGCCATCGATCGCGGACGGTAC-3'
	Clai

The *Bam*HI-*Sac*I fragment (0.38 kb) encoding the mutated residues in pSelect-BS plasmid derivatives was excised and used to replace the corresponding segment on pMUS1 to generate a series of mutant derivatives of pMUS1 (pMUS-IR6LE, pMUS-GI19EL, pMUS-G19N, pMUS-P30L and pMUS-YA33SM). Each 0.93 kb *Pst*I-*Sac*I fragment on these pMUS derivatives was isolated and ligated with the 5.4 kb *Pst*I-*Sac*I fragment of pMUL1 to generate the mutant derivatives of pSP105 (pSP105-IR6LE, pSP105-GI19EL, pSP105-G19N, pSP105-P30L and pSP105-YA33SM) carrying the designed mutations in the signal peptide of Npr (Figure 2).

Deletion mutagenesis of the mature region of *npr*

To construct pSP105ΔC, the 81 bp *Sca*I fragment at the 3' end of the *npr* gene was deleted, resulting in an excision of the 27 residues at the C-terminus of the Npr mature domain (Figure 3). To construct pSP105ΔM, an *Xba*I amber linker (BioLab) was introduced into the unique *Sma*I restriction site of plasmid pVU1200 [13], and the 1.2 kb *Pvu*II fragment of the resulting plasmid pVU1200ΔM was recovered and ligated with the 5.1 kb *Pvu*II DNA fragment of pVU5100 [13]. In pSP105ΔM the expected Npr derivative contained 284 residues of mature domain in addition to the signal sequence and propeptide region, followed

by an additional four residues (Pro-Ser-Leu-Asp) from the linker region at the C-terminus of the protein (Figure 3). To construct pSP105ΔD, plasmid pVU1200 was treated with restriction enzyme *Dra*III and repaired by T4 DNA polymerase. The 1.02 kb *Pvu*II fragment of the resulting plasmid pVU1200ΔD was excised and re-ligated with the 5.1 kb *Pvu*II fragment of plasmid pVU5100 to produce plasmid pSP105ΔD. This construct was expected to produce the fusion protein containing the Npr mature domain's N-terminal 42 residues, followed by an out-of-frame polypeptide sequence of 162 residues (Figure 3).

Construction of mutated secretory expression plasmids for propeptide

To construct the plasmids for directing the propeptide secretion by the mutated signal peptides, a similar strategy as described previously for the construction of secretory plasmid for the propeptide region was used [14]. Briefly, the 0.93 kb *Sac*I-*Pst*I fragment of each corresponding signal peptide mutant derivative of pMUS (pMUS-IR6LE, pMUS-GI19EL, pMUS-G19N and pMUS-P30) was ligated to the 5.4 kb *Sac*I-*Pst*I fragment of pSP105-Pro [14], resulting in pSP6LE-Pro, pSP19EL-Pro, pSP19N-Pro and pSP30L-Pro.

Preparation of the antiserum against pro-Npr

To overexpress the propeptide in *E. coli* for use in the antigenic generation of antiserum, the expression plasmid pET25b-Pro was constructed. The DNA segment that encodes for the pro-Npr region was amplified by PCR with the primers PS (5'-CGTACCGTCCGCCTCCATGGCCGCCCCCGCCCCG-3') and END (5'-CTCGGTACCCGGGCTAGCGTACTGGATG-GTGTGG-3'). The amplified fragment encompassed the complete region of pro-Npr and also contained *Nco*I and *Nhe*I sites at the 5' and 3' ends of the fragment respectively. This 1.5 kb PCR fragment was cloned into the *Sma*I site of pUC19 and re-excised as an *Nco*I-*Nhe*I fragment and inserted into the corresponding sites of 5.5 kb *E. coli* expression vector pET25b(+) (Novagen Co.) to generate the plasmid pET25b(+)-proNpr (His). Then the *Sac*I-*Kpn*I (0.74 kb) fragment of pMUL-Pro [14] was used to replace the corresponding fragment in pET25b(+)-proNpr (His) to produce plasmid pET25b-Pro. In this construct the translation stop codon was introduced into the last codon of the propeptide; therefore an *E. coli* strain harbouring this expression construct produces the propeptide directed by the pelB leader sequence after induction with isopropyl β-D-thiogalactoside [23]. *E. coli* strain BL21 harbouring this expression construct was cultured overnight in Luria-Bertani medium and induced for 3 h at 37 °C with 1 mM isopropyl β-D-thiogalactoside before harvest. The expected pelB-propeptide fusion protein (molecular mass 24 kDa) was separated on SDS/PAGE, and the protein band was sliced, eluted and used as antigen for the immunization of rabbits.

Purification of Npr from signal peptide mutants, and assay of Npr activity

The mature Npr was purified from the culture medium of *S. lividans* TK64 harbouring various pSP105 derivatives essentially as previously described [12]. Briefly, *S. lividans* were cultured in R2YE/starch liquid medium for 24–36 h [13]. The acetone-concentrated pellets of crude supernatants of media were dissolved in 50 mM sodium citrate buffer, pH 5.2 (buffer A), and then applied to a Fractogel CM TSK650(s) ion-exchange column. After washing with 130 mM NaCl in buffer A, the bound Npr was eluted by 0.15 M NaCl in buffer A.

THIS PAGE BLANK (USPTO)

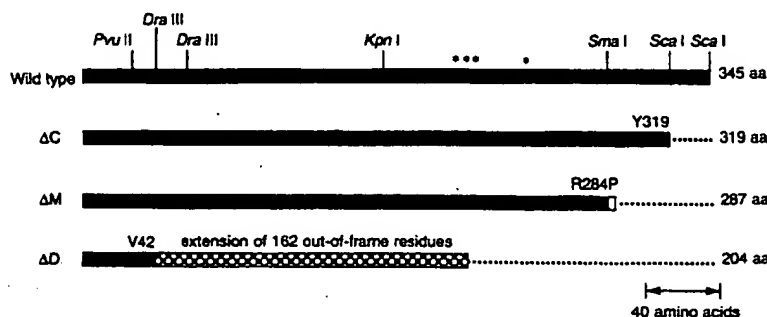


Figure 3 Schematic representation of deletions in the mature domain of Npr

The solid line at the top represents the coding sequence of the wild-type mature Npr. The asterisks indicate the regions coding for the active sites of Npr [13]. In the three deletion mutants shown, the dotted line represents the deleted portion of Npr mature region and the chequered bar indicates the frame-shifted region after position 42. The open bar in ΔM indicates the commercial amber linker region. The amino acid numbering is relative to the N-terminus (+1) of mature Npr. The sizes of the truncated Npr forms are indicated at the right (aa, amino acids). The nature and position of the residues at the junction of deletion or frameshift in each mutant are also indicated.

The phenotype of Npr was examined by the clear zones on R2YE agar [24] containing 2% (w/v) casein [13]. The protease level in liquid tryptic soy broth (TSB, Difco) culture supernatant was measured by using Azocoll (Sigma) as substrate [12,25]. One unit of protease activity was defined as previously reported [25].

Western blotting

Rabbit antibodies against Npr (35 kDa) and pro-Npr (57 kDa) were prepared as described elsewhere [13]. Antiserum against the propeptide of Npr (Pro) was prepared as described above. Proteins separated by SDS/PAGE were electroblotted on a nitrocellulose filter (Schleicher and Schuell) with a Bio-Rad mini trans-blot apparatus. Antigens were detected by anti-Npr, anti-pro-Npr [13] or anti-Pro antibodies. Goat anti-rabbit IgG conjugated with horseradish peroxidase (Bio-Rad) was used as the secondary antibody, and chromagenic development was performed as previously described [14].

RESULTS

Altered signal peptide sequence affects the secretion of Npr

The N-terminal 34 residues of prepro-Npr exhibited a typical prokaryotic signal peptide: three positive charged residues (Arg-5, Arg-7 and Lys-10) followed by an 11-residue hydrophobic stretch (Ala-12 to Ala-22) and a small side-chain amino acid at the putative signal peptidase cleavage site (Ala-34 and Ala-35) [12,26] (see also Figure 2). To elucidate whether mutations in the signal sequence would affect the secretion and maturation of Npr, five individual mutations were generated in the signal peptide (Figure 2; see the Experimental section for construction of mutants). In IR6LE the net charge on the N-terminal basic region was decreased from +3 to +1. In GI19EL and G19N a negatively charged (Glu) or a neutral polar (Asn) residue was incorporated into the central portion of the hydrophobic region to shorten the hydrophobic length. In P30L the proline residue that supposedly breaks the hydrophobic core α -helix structure [26] was replaced by leucine. In YA33SM the Tyr-Ala dipeptide upstream from the processing site was altered to Ser-Met.

When examined for the Npr activity on casein plates, all mutants displayed positive results (results not shown). However, the use of Azocoll as the assay substrate for secreted Npr in liquid broth revealed different profiles of protease activity among these mutants and the wild-type (Figure 4). Although the onset of extracellular protease production in all strains examined

exhibited a strong correlation with cell growth, the wild-type strain secreted Npr much earlier (12–18 h culture). The wild-type strain reached the peak of extracellular Npr activity at 24 h of

culturing time and declined thereafter. All mutants except GI19EL reached their peaks of Npr activity after 36–48 h of culturing and the duration of the expression was sustained from 18 to 24 h, which was similar to the wild-type strain. All mutants exhibited a decrease in secretory Npr activity. In mutants IR6LE and YA33SM the secreted Npr activity was 54–55% of the wild-type; in mutants G19N and P30L the secreted Npr activity was 33% and 18% of the wild-type respectively. In mutant GI19EL the Npr activity was less than 10% of the wild-type.

When analysed by immunoblot with anti-pro-Npr antiserum, the amounts of secreted Npr protein (P35) from the mutants correlated well with their activities (see Figures 4 and 5). A decrease in the secreted Npr to between one-third and one-half was found in IR6LE, YA33SM and G19N, and a decrease in P30L and GI19EL to between one-fifth and one-tenth, compared with the wild-type strain. Interestingly, a distinct band estimated at about 57 kDa, which accounted for 30–50% of the Npr species, was detected in mutants GI19EL and P30L at 12–54 h (mutant GI19EL) or 36–54 h (mutant P30L) of culturing (Figure 4). A similar protein species with weaker intensity was also found in 12–24 h cultures of the wild-type strain and much later in mutants IR6LE (36–48 h), G19N (24–30 h) and YA33SM (30–54 h). This protein species presumably represented the unprocessed pro-Npr species (p57) [13]. Several smaller protein species of 13–21 kDa, which probably represented the degradative forms of the Npr, were also detected in the immunoblots of the wild-type and mutant strains.

Taking these results together, we concluded that the integrity of the signal peptide is important for the export of Npr and its derivatives. Alterations within the four major regions of the signal peptide attenuated the export of Npr protein. In some mutants (GI19EL and P30L) the lesions seemed to delay the proteolytic processing of the secreted Npr precursor, resulting in the accumulation of unprocessed intermediate species.

A certain degree of variability in cell growth pattern (as reflected by the measurement of the total cellular proteins) was present in these signal peptide mutants compared with the wild-type strain (Figure 4). This suggests that the export defect in Npr might cause perturbations in cellular physiology. A similar observation was found in a signal peptide mutant of the major

THIS PAGE BLANK (USPTO)

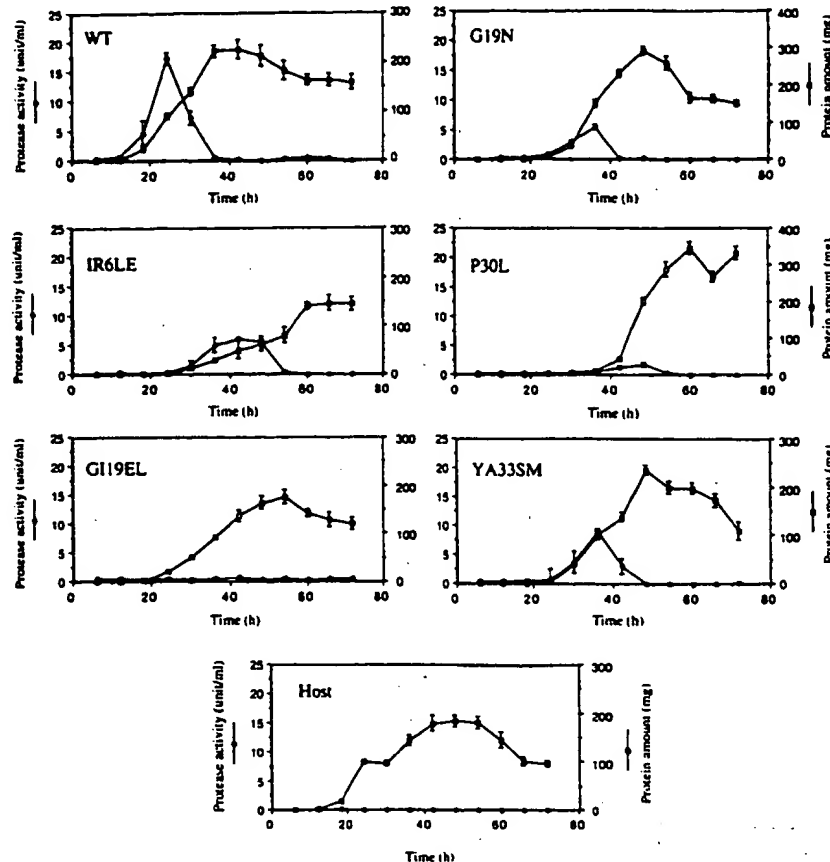


Figure 4 Analysis of protease activity in the strains harbouring the signal peptide mutations

TK64 harbouring wild-type or the *npr* signal peptide mutants were cultured in TSB medium for 8–60 h. The secreted Npr activity was determined from the culture supernatant as described in the Experimental section. The total protein in 50 ml of culture was also determined [67]. The results shown in the Figure are means \pm S.E.M. for three independent experiments performed in duplicate. Abbreviations: Host, TK64; WT, TK64[pSP105]; IR6LE, TK64[pSP105-IR6LE]; GI19EL, TK64[pSP105-GI19EL]; G19N, TK64[pSP105-G19N]; P30L, TK64[pSP105-P30L]; YA33SM, TK64[pSP105-YA33SM].

outer membrane lipoprotein of *E. coli* [27], where a deletion mutation in the lipoprotein signal peptide can cause severe effects on the viability of the host cell, resulting in cessation of growth and rapid cell death. Interestingly, in Npr this degree of interference with cell growth was not correlated with the severity of the export defect exerted by the respective signal peptide mutation introduced. The exact reason for this phenomenon is not clear, especially when one considers that mutant IR6LE showed a significantly decreased cell growth in spite of its exhibiting only a moderate inhibition of Npr production, whereas mutants GI19EL and P30L displayed a strong inhibition of Npr production but their cell growth showed a slight enhancement (Figure 4). Presumably this indicates that, in addition to the differences noted in the secretion of Npr, the pleiotropic effect of the Npr signal peptide mutation on the host cell might differ between mutants.

Npr derived from the signal peptide mutants has the same specific activity as that from the wild-type

To investigate whether mutations introduced into the signal peptide sequence might affect the folding of prepro-Npr and thus decrease the protease activities, the specific activities of Npr from the mutants were determined (Table 1). The results indicated that with the possible exception of G19N the specific activities of the

Npr from the other mutants (IR6LE and YA33SM) were essentially the same as that of the wild-type (Table 1). Results were not available for GI19EL and P30L, owing to the difficulty encountered in purifying the low quantities of Npr produced by these two mutants. Nevertheless the fact that similar levels of decrease in Npr export were found by Western blot and by enzymic assay strongly suggests that the same conclusion might be applicable to these two mutants.

Defects in the signal peptide sequence affect the secretion of propeptide

The propeptide of Npr can be secreted when fused directly to the signal peptide in the absence of the Npr mature domain [14]. To establish whether the mutations in signal sequence affect the secretion of propeptide as it did the secretion of Npr, the propeptide sequence was fused to the four mutant signal peptide sequences (IR6LE, GI19EL, G19N and P30L). Western blot analysis with anti-Pro antiserum demonstrated that all except P30L directed the secretion of the full-length (p22) propeptide into the medium at 16–20 h of culturing (Figure 6). However, the amounts of propeptide secreted by these mutants were much smaller than that secreted by the wild-type. The amounts of secreted propeptide directed by IR6LE, G19N, GI19EL and P30L at 16 or 20 h of culturing were approx. 88%, 68%, 23%

THIS PAGE BLANK (USPTO)

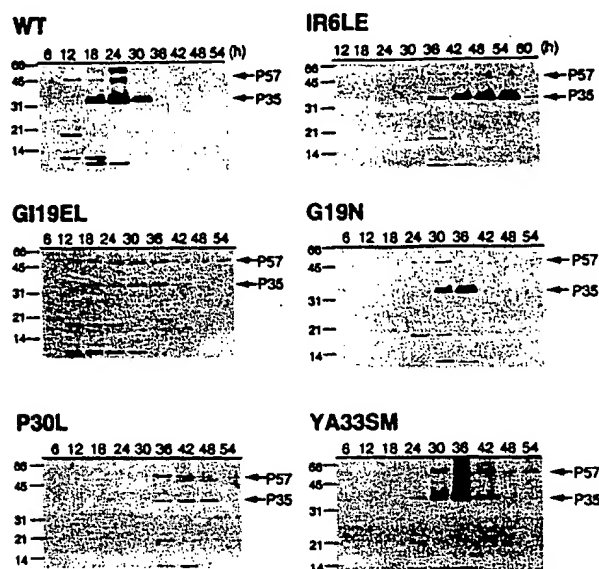


Figure 5 Western blot analysis of Npr production in Npr signal peptide mutant strains

The strains were cultured in TSB medium for the indicated times (6–60 h), and the culture supernatant (300 μ l) was precipitated with 10% (w/v) trichloroacetic acid at 4 °C. The recovered pellets were dissolved in sample buffer and analysed by SDS/PAGE [12.5% (w/v) gel] [68], followed by Western blot analysis using anti-pro-Npr (P57) antiserum as described in the Experimental section. Arrows indicate the positions of pro-Npr (P57) or mature Npr (P35). The designations of the host or *npr* strains are the same as those in Figure 4.

Table 1 Specific activity of purified Npr proteases from wild-type and various signal peptide mutants

The Npr proteases from wild-type and mutant strains were purified by column chromatography as described in the Experimental section. The protease activity was assayed by the method of Chavira et al. [25]. The results are means \pm S.E.M. for three independent experiments.

Enzyme	Specific activity (units/mg of protein)
Wild-type	623 \pm 121
IR6LE	573 \pm 21
GI19N	470 \pm 18
YA33SM	583 \pm 77

and 0% of that of the wild-type. Smaller protein species, which presumably represented the degraded propeptide, were also found. Thus it seems that those mutants with strong defects in the secretion of Npr (e.g. mutants GI19EL and P30L) were also defective in propeptide secretion (compare Figures 5 and 6).

Deletions of the mature domain abolish the maturation but not the secretion of Npr derivatives

To investigate whether the integrity of the mature region is essential for the maturation and secretion of Npr, three deletion mutations (Δ C, Δ M and Δ D) were created in the Npr mature region (see Figure 3). In Δ C the deletions spanned the C-terminal 26 residues (320–345). In Δ M the C-terminal 62 residues (284–345) were deleted and the C-terminus of the resulting mutant protein was replaced by four residues (Pro-Ser-Leu-Asp) encoded by the *Xba*I amber linker. In Δ D the N-terminal 42-residue fragment of the mature domain was fused with a different

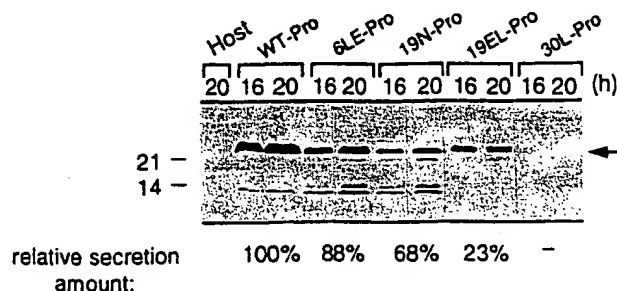


Figure 6 Western blot analysis of the propeptide of Npr directed by mutant signal peptides

Culture supernatants (600 μ l; 16 or 20 h) from host or different propeptide expression variants was precipitated with 10% (w/v) trichloroacetic acid at 4 °C. The recovered pellet was dissolved in sample buffer and analysed by SDS/PAGE [68], followed by Western blot analysis with anti-Pro antiserum. The arrow indicates the propeptide-specific bands. The band intensity was quantified by a densitometer and expressed as the relative amount to that of wild-type. Abbreviations: Host, TK64; WT-Pro, TK64[pSP105-Pro]; 6LE-Pro, TK64[pSP6LE-Pro]; 19N-Pro, TK64[pSP19N-Pro]; 19EL-Pro, TK64[pSP19EL-Pro]; 30L-Pro, TK64[pSP30L-Pro].

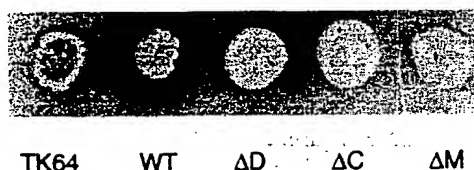


Figure 7 Analysis of Npr activity in Npr mature-region deletion mutants

Npr activity in the signal peptide mutants was revealed on the casein plate as described in the Experimental section. Abbreviations: TK64, host; WT, TK64[pSP105]; Δ D, TK64[pSP105 Δ D]; Δ C, TK64[pSP105 Δ C]; Δ M, TK64[pSP105 Δ M].

Table 2 Analysis of Npr activity in Npr mature-region mutants

The Npr proteases from wild-type and Npr mature-region mutant strains were assayed as described in the Experimental section. Protease activity is expressed as units/ml of culture supernatant. The designations for the host or *npr* strains are the same as in Figure 7. The data are means \pm S.E.M. for three independent experiments.

Culturing time (h)	Protease activity (units/ml of culture supernatant)				
	TK64	WT	Δ C	Δ D	Δ M
16	0.6 \pm 0.0	8.0 \pm 0.1	0.2 \pm 0.0	0.4 \pm 0.0	0.3 \pm 0.0
21	0.7 \pm 0.0	13.0 \pm 2.0	0.4 \pm 0.0	0.4 \pm 0.1	0.5 \pm 0.1
26	0.6 \pm 0.0	13.4 \pm 1.0	0.4 \pm 0.0	0.4 \pm 0.1	0.4 \pm 0.1
31	0.2 \pm 0.0	4.9 \pm 1.1	0.4 \pm 0.0	0.4 \pm 0.1	0.2 \pm 0.0

reading frame polypeptide sequence with 162 residues resulting from a frame shift beyond Val-42 of the mature region. The expected molecular masses of the truncated Npr protein secreted from Δ C, Δ M and Δ D were 34, 32 and 22 kDa respectively, if the mutant propeptides were processed as in the wild-type.

As shown in Figure 7, all three mutants displayed an Npr phenotype on casein plates, although the deletions in mutants Δ C and Δ M did not encompass the residues essential for the Npr activity (i.e. His-202, Glu-203, His-206 and Glu-240) [13]. No detectable protease activities were found in the liquid broth of these mutant strains either (Table 2). These results suggest that deletions in the mature domain of Npr might affect the matur-

THIS PAGE BLANK (USPTO)

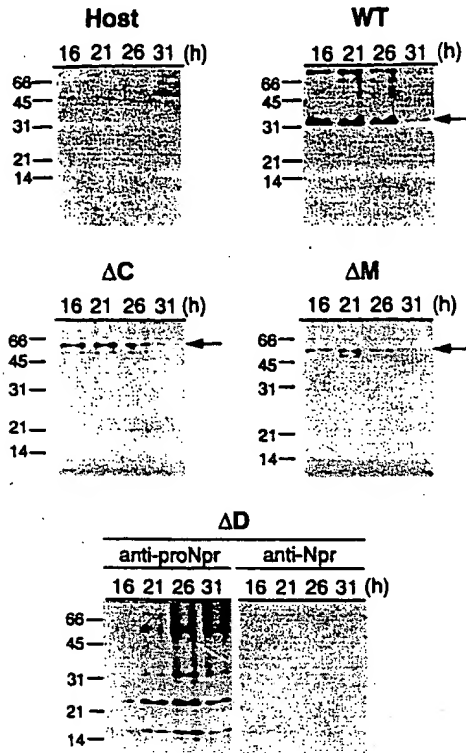


Figure 8 Western blot analysis of the production of Npr derivatives from the culture supernatants of *npr* mature-region mutants

The host and *npr* mature-region deletion mutants were cultured in TSB medium for 16–31 h, and the culture supernatant (600 μ l for host strain and 100 μ l for the wild-type or mutant strains) was precipitated with 10% trichloroacetic acid and the recovered pellets were dissolved in sample buffer and analysed by SDS/PAGE [68], followed by Western blot analysis with anti-pro-Npr antiserum. The immunoreactive bands specific to anti-pro-Npr antisera are indicated with arrows. The abbreviations for the wild-type or mutant strains are as in Figure 7.

ation and/or processing of Npr. Immunoblot analysis with the anti-pro-Npr antibody detected 56 and 54 kDa proteins secreted by ΔC and ΔM (Figure 8). The size range of these protein species corresponded to the expected truncated pro-Npr derivatives if the pro-Npr derivatives were not processed. In the frame-shift mutant, ΔD , several protein species (55, 34 and 24 kDa) exceeding the expected size (22 kDa) for the processed Npr fusion protein were detected by anti-pro-Npr antiserum but not by anti-Npr antiserum (Figure 8). These protein species were presumably the unprocessed Npr fusion protein derivatives. All these results indicate that a truncation as small as 26 residues of the C-terminus of the mature domain completely abolishes the auto-processing activity of Npr. Interestingly, these truncations did not block the secretion of these Npr derivatives.

DISCUSSION

In this report we have shown that the secretion of active Npr in *Streptomyces* is dependent on the presence of the intact N-terminal signal peptide. Substitution mutations in four major regions (basic region, hydrophobic core, β -turn region and signal-peptidase cleavage site) of the signal peptide sequence led to an attenuation of Npr export, both kinetically and quantitatively, although the degree of impairment varied (Figures 4 and 5). Similar results were obtained when these mutant signal peptides were used to direct the secretion of Npr propeptide in the absence

of the mature domain (Figure 6). Thus the integrity of the Npr signal peptide is important in the secretion of the Npr and propeptide. Compared with the signal peptides of other bacterial secretory proteins, several features of the Npr signal sequence are noted. For instance, although the translocation of a number of *E. coli* secretory proteins (e.g. lipoprotein, staphylokinase and maltose-binding protein) across the cytoplasmic membrane does not absolutely require a net positive charge at their N-terminus [28–30], we found that in mutant IR6LE even the retention of a net charge of +1 at the N-terminus still caused a negative effect on Npr export (see IR6LE in Figures 4 and 5). The requirement for charged residues for the secretion of Npr in *Streptomyces* may be due to the nature of the host membrane, or it could be unique to Npr.

The introduction of charged or polar residues to the hydrophobic core of the Npr signal peptide impaired the secretion of Npr and pro-Npr (see G119EL and G19N in Figures 4–6). Similar results were obtained with *E. coli* secretory proteins, although the severity of impairment varied considerably in different systems [2]. For instance, Goldstein et al. [31] introduced an asparagine residue (mutants I8N and V10N) into the hydrophobic core region of OmpA protein, which resulted in the complete blocking of secretion. In the Gly-14 to Asp-14 mutation (mutant 14D) in the lipoprotein signal sequence, translocation of precursor protein across the cytoplasmic membrane was not affected [32]. In contrast, the introduction of an aspartate or glutamate residue at positions 14 and 15 respectively of the LamB signal sequence markedly decreased secretion [33]. In our study, replacement of central-core Gly-Ile residues at positions 19 and 20 by Glu-Leu residues in mutant G119EL resulted in a severe impairment of Npr export and proteolytic processing (Figures 4 and 5). Less severe defects were conferred by the G19N mutation, in which the polar residue Asn was introduced (Figures 4 and 5). All these findings indicate that the N-terminus and hydrophobic region in the Npr signal peptide are sensitive to the lesions introduced.

It was also surprising to find that the alteration introduced in the residues preceding the signal peptidase cleavage site in the YA33SM mutant caused a moderate export defect in Npr (Figures 4 and 5), whereas a similar substitution at the corresponding region in MelC1 signal peptide yielded no detectable effects on export [8]. This contrast might reflect their respective secondary structures at the processing site. In addition we observed an almost completely defective signal peptide in mutant P30L, in which a mutation at the hydrophobic core–cleavage site junction supposedly disrupted the predicted β -turn (Figures 4 and 5). This finding agrees with the hypothesis that a β -turn structure four to six residues upstream of the cleavage site is required for precursor protein processing [26,34,35]. However, this result stands in contrast with the observation of Borchert and Nagarajan [7] that the export of *Bacillus amyloliquefaciens* levansucrase remained unaffected when its β -turn structure at the end of the hydrophobic core of the signal peptide was disrupted. It is likely that the sequence requirement in the Npr signal peptide is more constrained than other secretory proteins discussed here.

Several roles of the signal peptide in facilitating protein export have been proposed [1,36–38]. In the present study we cannot extrapolate the exact mechanisms by which these signal peptide mutants hinder or delay Npr export across the membrane. The mutations introduced into the signal peptide that deteriorate protein secretion might result from the alteration of the conformation or hydrophobicity of the signal peptide, which might affect its insertion into the membrane or its interactions with secretory components and/or alter the export competence of the secretory proteins attached to it [1,36–38]. Notably, when we

THIS PAGE BLANK (USPTO)

Table 3 The β -turn structure features of bacterial signal peptide sequences in secretory proteins predicted by the method of Chou and Fasman [39]

The window width for the analysis was four amino acid residues. The program used in this analysis was PC/GENE of Intelligenetics.

Strain	Protein	Signal peptide sequence	β -turn potential (residue number)	Reference
<i>S. cerevisiae</i> A3(2)	Agarase	MVNRDLIKWSAVALGAGAGLAGPAPAAHA	3 (31)	[50]
<i>S. griseus</i>	Amylase	MARRLATASLAVLAAAATALTAPTAAA	1 (28)	[51]
<i>S. hygroscopicus</i>	Amylase	MQQSRVLGGTLAGTVAAAAATVAPWPSQA	4 (30)	[52]
<i>S. venezuelae</i>	Amylase	MARKTVAAALALVAGAAVAVTGNAPQA	1 (28)	[53]
<i>S. limosus</i>	Amylase	MARRLATASLAVLAAAATALTAPTAAA	1 (28)	[54]
<i>S. cacaoi</i>	Npr	MPMFRIPLPKPAAIAAGGIGACIATVAVPSAYA	3 (34)	[13]
<i>S. albus</i>	β -Lactamase	VHPSTSRPSRRTLLTAGAALAAATLVPGTAHASSGGR	6 (39)	[55]
<i>S. sp36</i>	Xylanase	MNQDGKRYESEQNPPFSGLSRRGLVYGAGAGAAAVAGSGLLLPGTAHA	11 (50)	[56]
<i>S. antibioticus</i>	MelC1	MPELTRRRALGAAAVVAGVPLVALPAARA	0 (30)	[57]
<i>B. amyloliquefaciens</i>	Barnase	MKKRLSWISVCLLVLSAAGMLFSTA	0 (26)	[47]
<i>Bacillus</i>	Alkaline protease	MRGKKVWISLLFALALIFTMAFGSTSSQA	1 (30)	[47]
<i>Bacillus</i>	Neutral protease	MGLGKLSVAVAAFSMLTISLPGVQA	1 (27)	[47]
<i>B. amyloliquefaciens</i>	Amylase	MIQKRKRTVSFRLVLMCTLLFVSLPTKTS	0 (31)	[58]
<i>B. subtilis</i>	Amylase	MFAKRFKTSLLPLFAGFLLFLVLVAGPAA	0 (31)	[59]
<i>B. amyloliquefaciens</i>	Levansucrase	MNIKKIVKQATVLTFTTAGGATQDA	1 (28)	[7]
<i>E. coli</i>	LamB	MMITLRKLPLAVAVAAGVMSAQAMA	0 (25)	[60]
<i>E. coli</i>	Alkaline phosphatase	MKOSTIALALLPLFTPTKA	0 (21)	[61]
<i>E. coli</i>	OmpA	MKKTAIAIVALAGFATVQA	0 (25)	[62]
<i>E. coli</i>	Lipoprotein	MKATKLVLGAVILGSTLLAG	0 (20)	[63]
<i>E. coli</i>	Maltose-binding protein	MKIKTGARILALSALTMMASALA	0 (25)	[64]
<i>E. coli</i>	Staphylokinase	MLKRSLLFLTVLLLFSSSITNOVSA	0 (27)	[65]
<i>E. coli</i>	β -Lactamase	MSIQHFRVALUPFFAAFCPLVFA	1 (23)	[66]

used the Chou-Fasman [39] rule to predict the propensity of Npr signal peptide for β -turns, there are potentials for β -turns located at residues Leu-8, Ala-16 and Val-29. Moreover, for those mutants (G119EL and P30L) yielding severely decreased Npr secretion, their predicted β -turn structures within the Npr signal peptide had altered. Interestingly, a similar observation can be made on the mutations of neutral protease signal peptide from *B. amyloliquefaciens* when fused with human parathyroid hormone, where changes in the predicted β -turn structure near the signal peptidase cleavage site substantially affect the secretory efficiency of the signal peptide [40]. Unexpectedly, an analysis of several secretory proteins from *Streptomyces* also showed a higher propensity for β -turn structure in their signal peptide regions than those from *Bacillus* or *E. coli* (Table 3). However, no such feature was observed in the signal peptide of MelC1 protein. This may account for the differential sensitivity to the similar mutation introduced into the cleavage site of the MelC1 or Npr signal peptide, as described earlier. The β -turn structure might be one of the requirements for proper function of the Npr signal peptide. However, the generalization of the β -turn requirement for secretion of other *Streptomyces* secretory proteins must await additional information obtained from other *Streptomyces* signal peptide mutants.

The mature region has been shown to have a profound impact on the export function of several secretory proteins [41–47]. Altman et al. [44], for example, have demonstrated that both the signal peptide and mature regions of LamB protein are required for the interaction with the export factor SecB in *E. coli*. In addition, the charge distribution as well as the secondary structure of the N-terminal mature region adjacent to the signal peptide have been recognized as important determinants for the compatibility of protein secretion [35,41–43,45]. In the present study, neither the N-terminal frame-shift nor the C-terminal deletion of Npr mature domain affected the secretion of Npr derivative (Figure 8). It had previously been observed that, although the integrity of propeptide is not essential for the export of the pro-

Npr derivatives, removal of propeptide while preserving intact mature region completely abolished secretion [14]. Thus, in contrast with the propeptide, if the mature region is directly fused with the signal peptide it remains incompetent for secretion. This dependence on the propeptide sequence for secretion might be due to structural incompatibility between the C-terminus of the signal peptide and the N-terminus of the mature domain, or possibly to the fact that signal peptide alone, without the help of the propeptide, cannot maintain the Npr mature region in an export-competent folded state. The presence of two positively charged residues at the N-terminus of the Npr mature sequence (positions 4 and 5), which has been suggested to be detrimental to protein secretion [35,42–45], is consistent with the first hypothesis. Nevertheless the second view is also very likely, because one of the major functions of the propeptide in protease is to act as an intramolecular chaperone for the modulation of protein conformation [48,49].

In summary, on the basis of our previous work [13,14] and results presented here it is clearer that in the processes of Npr secretion and maturation, the three distinct domains of Npr are essential and confer different functions. Although a functional signal peptide is essential for secretion, the propeptide guides the precursor domains to achieve a conformation that is competent for the export and maturation of Npr. Finally, the properly folded mature region yields the Npr activity and removes the propeptide sequence. Therefore the secretory mechanism of Npr precursor in *Streptomyces* is fundamentally analogous to the transport of precursor protein across the cytoplasmic membrane of Gram-negative bacteria in that both are signal-peptide-dependent and require a transport-competent conformation of precursor protein.

We thank Dr. C.W. Chen, Dr. S.J. Lo and Dr. C. Weaver for comments and critical reading of the manuscript, and Dr. L.-L. Liaw for help in manuscript preparation. This work was supported by Grant NSC 83-0412-B010-004 from the National Science Council to Y.-H.W.L.

THIS PAGE BLANK (USPTO)

REFERENCES

- 1 Gierasch, L. M. (1989) *Biochemistry* **28**, 923–930
- 2 Gennity, J., Goldstein, J. and Inouye, M. (1990) *J. Bioenerg. Biomembr.* **22**, 233–269
- 3 Model, P. and Russel, M. (1990) *Cell* **61**, 739–741
- 4 Izard, J. W. and Kendall, D. A. (1994) *Mol. Microbiol.* **13**, 765–773
- 5 Palva, I., Sarvas, M., Lehtovaara, P., Sibakov, M. and Kaariainen, L. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 5582–5586
- 6 Nakamura, K., Fujita, Y., Itoh, Y. and Yamane, K. (1989) *Mol. Gen. Genet.* **216**, 1–9
- 7 Borchert, T. V. and Nagarajan, V. (1991) *J. Bacteriol.* **173**, 276–282
- 8 Leu, W.-M., Chen, L.-Y., Liaw, L.-L. and Lee, Y.-H. W. (1992) *J. Biol. Chem.* **267**, 20108–20113
- 9 Lee, Y.-H. W., Chen, B.-F., Wu, S.-Y., Leu, W.-M., Lin, J.-J., Chen, C. W. and Lo, S. J. (1988) *Gene* **65**, 71–81
- 10 Chen, L.-Y., Leu, W.-M., Wang, K.-T. and Lee, Y.-H. W. (1992) *J. Biol. Chem.* **267**, 20100–20107
- 11 Chen, L.-Y., Chen, M.-Y., Leu, W.-M., Tsai, T.-Y. and Lee, Y.-H. W. (1993) *J. Biol. Chem.* **268**, 18710–18716
- 12 Chang, P.-C., Kuo, T.-C., Tsugita, A. and Lee, Y.-H. W. (1990) *Gene* **88**, 87–95
- 13 Chang, P.-C. and Lee, Y.-H. W. (1992) *J. Biol. Chem.* **267**, 3952–3958
- 14 Chang, S.-C., Chang, P.-C. and Lee, Y.-H. W. (1994) *J. Biol. Chem.* **269**, 3548–3554
- 15 Ikemura, H., Takagi, H. and Inouye, M. (1987) *J. Biol. Chem.* **262**, 7859–7864
- 16 Zhu, X., Ohla, Y., Jordan, F. and Inouye, M. (1989) *Nature (London)* **339**, 483–484
- 17 Wetmore, D. R., Wong, S.-L. and Roche, R. S. (1992) *Mol. Microbiol.* **12**, 1593–1604
- 18 Randall, L. L. and Hardy, S. J. S. (1986) *Cell* **46**, 921–928
- 19 Pugsley, A. P. (1993) *Microbiol. Rev.* **57**, 50–108
- 20 Hopwood, D. A., Kieser, T., Wright, H. M. and Bibb, M. J. (1983) *J. Gen. Microbiol.* **129**, 2257–2269
- 21 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 22 Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Kieser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M. and Schrempf, H. (1985) *Genetic Manipulation of Streptomyces: A Laboratory Manual*, John Innes Foundation, Norwich
- 23 Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60–89
- 24 Chater, K. F., Hopwood, D. A., Kieser, T. and Thompson, C. J. (1982) *Curr. Top. Microbiol. Immunol.* **96**, 69–95
- 25 Chavira, R. J., Burnett, T. J. and Hageman, J. H. (1984) *Anal. Biochem.* **136**, 446–450
- 26 Perlman, D. and Halvorson, H. O. (1983) *J. Mol. Biol.* **167**, 392–409
- 27 Pollitt, N. S. and Inouye, M. (1988) *J. Bacteriol.* **170**, 2051–2055
- 28 Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K. and Inouye, M. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3438–3441
- 29 Iino, T., Takahashi, M. and Sako, T. (1987) *J. Biol. Chem.* **262**, 7412–7417
- 30 Puziss, J. W., Fikes, J. D. and Bassford, Jr., P. J. (1989) *J. Bacteriol.* **171**, 2303–2311
- 31 Goldstein, J., Lehnhardt, S. and Inouye, M. (1991) *J. Biol. Chem.* **266**, 14413–14417
- 32 Sung, C. Y., Gennity, J. M., Pollitt, N. S. and Inouye, M. (1992) *J. Biol. Chem.* **267**, 997–1000
- 33 Stader, J., Benson, S. A. and Silhavy, T. J. (1986) *J. Biol. Chem.* **261**, 15075–15080
- 34 Kuhn, A. and Wickner, W. (1985) *J. Biol. Chem.* **260**, 15914–15918
- 35 von Heijne, G. (1986) *J. Mol. Biol.* **192**, 287–290
- 36 Park, S., Liu, G., Topping, T. B., Cover, W. H. and Randall, L. (1988) *Science* **239**, 1033–1035
- 37 Randall, L. L. and Hardy, S. J. S. (1989) *Science* **243**, 1156–1159
- 38 de Vrije, G. J., Batenburg, A. M., Killian, J. A. and de Kruijff, B. (1990) *Mol. Microbiol.* **4**, 143–150
- 39 Chou, P. Y. and Fasman, G. D. (1979) *Biophys. J.* **26**, 367–384
- 40 Saunders, C. W., Pedroni, J. A. and Monahan, P. M. (1991) *Gene* **102**, 177–181
- 41 Duffaud, G. and Inouye, M. (1988) *J. Biol. Chem.* **263**, 10224–10228
- 42 Li, P., Beckwith, J. and Inouye, H. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7685–7689
- 43 Yamane, K. and Hizushima, S. (1988) *J. Biol. Chem.* **263**, 19690–19696
- 44 Altman, E., Emr, S. D. and Kumamoto, C. A. (1990) *J. Biol. Chem.* **265**, 18154–18160
- 45 MacIntyre, S., Eschbach, M.-L. and Mutschler, B. (1990) *Mol. Gen. Genet.* **221**, 466–474
- 46 Tanji, Y., Gennity, J., Pollitt, S. and Inouye, M. (1991) *J. Bacteriol.* **173**, 1997–2005
- 47 Chen, M. and Nagarajan, V. (1993) *Mol. Gen. Genet.* **239**, 409–415
- 48 Baker, D., Shiau, A. and Agard, D. A. (1993) *Curr. Opin. Cell Biol.* **5**, 966–970
- 49 Shinde, U. and Inouye, M. (1993) *Trends Biochem. Sci.* **18**, 442–446
- 50 Kendall, K. and Cullum, J. (1984) *Gene* **29**, 315–321
- 51 Vigal, T., Gil, J. A., Daza, A., Garcia-Gonzalez, M. D., Villadas, P. and Martin, J. F. (1991) *Mol. Gen. Genet.* **231**, 88–96
- 52 Hoshiko, S., Makabe, O., Nohji, C., Katsumata, K., Satoh, E. and Nagaoka, K. (1987) *J. Bacteriol.* **169**, 1029–1036
- 53 Virolle, M. J., Long, C. M., Chang, S. and Bibb, M. J. (1988) *Gene* **74**, 321–334
- 54 Long, C. M., Virolle, M. J., Chang, S. Y., Chang, S. and Bibb, M. J. (1987) *J. Bacteriol.* **169**, 5745–5754
- 55 Dehottay, P., Dusart, J., De Meester, F., Joris, B., Van Beeumen, J., Ericum, T., Frère, J. M. and Ghuyssen, J. M. (1987) *Eur. J. Biochem.* **166**, 345–350
- 56 Iwasaki, A., Kishido, H. and Okanishi, M. (1987) *J. Antibiot. (Tokyo)* **39**, 985–993
- 57 Berman, V., Filpula, D., Herber, W., Bibb, M. and Katz, E. (1985) *Gene* **37**, 101–110
- 58 Palva, I., Pettersson, R. F., Kalkkinen, N., Lehtovaara, P., Sarvas, M., Soderlund, H. and Kaariainen, L. (1981) *Gene* **15**, 43–51
- 59 Yang, M., Galizzi, A. and Henner, D. (1983) *Nucleic Acids Res.* **11**, 237–249
- 60 Emr, S. D., Schwartz, M. and Silhavy, T. J. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5802–5806
- 61 Inouye, H., Barnes, W. and Beckwith, J. (1982) *J. Bacteriol.* **149**, 434–439
- 62 Mowbray, N. R., Nakamura, K. and Inouye, M. (1980) *J. Biol. Chem.* **255**, 27–29
- 63 Nakamura, K. and Inouye, M. (1979) *Cell* **18**, 1109–1117
- 64 Bedouelle, H., Bassford, P. J., Fowler, A. V., Zabin, I., Beckwith, J. and Hoffnung, M. (1980) *Nature (London)* **285**, 78–81
- 65 Iida, A., Groarke, J. M., Park, S., Thom, J., Zabicky, J. H., Hazelbauer, G. L. and Randall, L. L. (1985) *EMBO J.* **4**, 1875–1880
- 66 Sutcliffe, J. G. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3737–3741
- 67 Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254
- 68 Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685

THIS PAGE BLANK (USPTO)

Von Willebrand factor storage requires intact prosequence cleavage site

Agnès M. Journet^{2)a}, Simin Saffaripour^a, Elisabeth Martin Cramer^b, Danièle Tenza^b,
Denisa D. Wagner^{1)a}

^a Center for Hemostasis and Thrombosis Research, Division of Hematology-Oncology, New England Medical Center and Department of Anatomy and Cellular Biology, Tufts University School of Medicine, Boston, MA/USA

^b Unité INSERM U348 and Institut des Vaisseaux et du Sang, Hôpital Lariboisière, Paris/France

Received October 12, 1992

Accepted November 6, 1992

von Willebrand factor—propeptide cleavage—regulated secretion—storage granules

Large multimers of the adhesive glycoprotein von Willebrand factor (vWf) are stored in endothelial cells in rod-shaped granules called Weibel-Palade bodies, while small multimers are secreted constitutively. Expression of pro-vWf in other cells with a regulated pathway of secretion, results in formation of vWf-containing storage granules that have a morphology similar to Weibel-Palade bodies. vWf expressed without its prosequence is not stored. To evaluate the importance of prosequence cleavage in vWf storage, the Arg at position -1, known to be necessary for cleavage, was mutated to Gly. Transfection of this cleavage mutant into two cell lines with a regulated pathway of secretion (RIN 5F and AtT-20 cells) led to the formation of large multimers. However, treatment of the cell lysates by the enzyme endoglycosidase H (Endo-H) did not reveal significant amounts of intracellular Endo-H-resistant vWf, which indicates the absence of a pool of stored processed vWf. In addition, no Weibel-Palade body-like structure was detected in these cells by immunofluorescence labeling with anti-vWf antiserum. Electron microscopy and immunocytochemistry of RIN 5F cells expressing the pro-vWf mutant confirmed the absence of Weibel-Palade body-like structures. In addition, anti-vWf-linked gold particles were found in the ER, occasionally in rounded granules and particularly in lysosomal structures which were abundant. We conclude that the formation of large aggregates is not sufficient to induce efficient vWf storage, and that the lack of cleavage

of the prosequence may direct the mutant pro-vWf molecule to a degradative pathway. Therefore, the prosequence cleavage is a requirement for vWf storage.

Introduction

Endothelial cells lining the vessel wall possess elongated storage granules called Weibel-Palade bodies [42] whose function is important in injury repair. These organelles are found only in endothelium and contain the soluble adhesive protein von Willebrand factor (vWf) and the transmembrane receptor for monocytes and neutrophils, named P-selectin (CD62) [1, 20, 37]. Stimulation of the endothelial cells leads to the secretion of large vWf multimers that are most potent in platelet plug formation (for review, see [34]). In vitro, small vWf multimers are also secreted through a constitutive pathway [28].

Both vWf [41] and P-selectin [6, 12] contain a generally recognized targeting signal that allows their storage in heterologous cells with a regulated pathway of secretion. In the case of vWf it was shown that the presence of a large prosequence with which the molecule is first synthesized [15, 35] is required for formation of vWf storage granules [41]. When the wild type pro-vWf is expressed in AtT-20 cells (mouse pituitary cell line) or in RIN 5F cells (rat insulin-secreting B cell line), vWf storage granules are formed that are morphologically very similar to Weibel-Palade bodies and distinct from the endogenous hormone granules. The Weibel-Palade body-like organelles do not form in cells transfected with vWf cDNA from which the prosequence was deleted [41].

Another important intracellular function of the prosequence is in the formation of vWf multimers from building blocks of dimeric molecules. Deletion of the prosequence results in the formation of only dimeric molecules in the transfected cells [32, 44]. The multimerization of pro-vWf dimers by disulfide bonds located in the N-terminal portion

Abbreviations. CMV Cytomegalovirus.—Endo-H Endoglycosidase H.—ER Endoplasmic reticulum.—HMW High molecular weight.—kb Kilobases.—SDS Sodium dodecyl sulfate.—vWf von Willebrand factor.

¹⁾ Dr. Denisa D. Wagner, Division of Hematology-Oncology, New England Medical Center, Box 832, 750 Washington Street, Boston, MA 02111/USA.

²⁾ New address: Unité INSERM U238, DBMS/ICH, Centre d'Etudes Nucléaires de Grenoble, F-38041 Grenoble Cédex/France.

THIS PAGE BLANK (USPTO)

of the mature subunit takes place in the acidic environment of the trans Golgi apparatus and/or in the secretory granules. This process is likely followed by the cleavage of the prosequence at a dibasic amino acid cleavage site (for review, see [19]). Inhibition of this cleavage by site-specific mutagenesis of the cleavage site, does not impair the vWf multimerization process in transfected cells [33, 44]. In the endothelial Weibel-Palade bodies, the free propolypeptide and the mature vWf subunits are stored in stoichiometric amounts but their relative location within the granule is unknown [7, 39].

In the present study, we are interested in determining if successful multimer formation is linked to vWf storage and if inhibition of the prosequence cleavage effects the formation of vWf storage granules. Therefore, we generated an uncleavable pro-vWf encoding cDNA and transfected it in cell lines with a regulated pathway of secretion.

Materials and methods

Reagents

All the restriction and modification enzymes were purchased from Bethesda Research Laboratories (Gaithersburg, MD/USA), or New England Biolabs (Boston, MA/USA). The BamHI-EcoRI adaptors were from New England Biolabs. The oligonucleotide-directed in vitro mutagenesis system was from Amersham International (Amersham, Bucks./UK). Protein A-Sepharose CL-4B was purchased from Sigma (St. Louis, MO/USA).

Antibodies

Rabbit polyclonal antiserum against vWf was purchased from American Bioproducts (Parsippany, NJ/USA) or Dakopatts (Carpinteria, CA/USA). The preparation of monoclonal antibodies against the propolypeptide was described [40]. Goat anti-rabbit immunoglobulin fraction coupled to 10 nm colloidal gold was from Amersham (Les Ulis/France).

Oligonucleotide synthesis

Synthetic oligonucleotides were prepared with an automated DNA synthesizer type 381A (Applied Biosystems, Foster City, CA/USA) by M. Jacobs (New England Medical Center, Boston, MA) and purified on an OPC column (Applied Biosystems), according to the manufacturer protocol, and dissolved in 1 mM EDTA in 10 mM Tris (pH 7.4) for further use.

Vectors and recombinant DNA construction

pCMV2. This vector, kindly provided by Dr. Richard Goodman (Portland, OR/USA), contains the cytomegalovirus (CMV) promoter and the SV40 small t antigen polyadenylation signal, on both sides of the BglII insertion site. This eukaryotic sequence is inserted into the ClaI site of the bacterial Bluescript vector (Stratagene, La Jolla, CA/USA).

Construction of pCMV4. pCMV2 was digested with EcoRI and the two resulting fragments, Bluescript and the eukaryotic sequence, were blunt-ended with the Klenow DNA polymerase and religated. Therefore, the resulting molecule had endogenous EcoRI sites removed so that they would not interfere with subsequent construction steps. BamHI-EcoRI adaptors were inserted in the BglII site to produce pCMV4. The newly created EcoRI site was used for insertion of cDNA in pCMV4.

Construction of pCMV4-pro-vWf Gly763. The pro-vWf 482 bp HindIII-BamHI fragment was subjected to oligonucleotide-directed mutagenesis, using the Amersham system. The chosen oligonucleotide (5' GCA GCA AAG* GGA GCC TAT 3', positions 2279 to 2296, with a substitution from A to G at position 2287, marked with an asterisk; position 1 is the first base of the initiator methionine [3]) gives the same mutation as the one used by Verweij et al. [33]: it transforms the Arg, at the -1 position with regard to the propolypeptide cleavage site, into a Gly. The vWf 2.5 kb-EcoRI-HindIII fragment, the mutated 482 bp HindIII-BamHI fragment and the vWf 6 kb BamHI-EcoRI fragment were purified. In a single reaction, these three fragments were inserted in pCMV4 linearized by EcoRI to produce pCMV4-pro-vWf Gly763. The newly made plasmid was analyzed by restriction analysis, and the A to G substitution was checked by sequencing the complete insert in the M13 derivative that resulted from the mutagenesis reaction.

Cell culture and transfection

AtT-20 cells and COS-1 cells were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum, and penicillin/streptomycin (100 U/ml) (Gibco Laboratories, Grand Island, NY/USA). RIN 5F was maintained in DMEM supplemented with 10% heat-inactivated horse serum, 0.2% glucose and penicillin/streptomycin (100 U/ml). COS-1 cells and AtT-20 cells were transfected by the calcium phosphate procedure [8, 43], using 40 µg of plasmid per 10 cm dish as described [41]. RIN 5F were transfected by the same procedure except that the cells were in suspension. After 15 min incubation with the calcium phosphate-DNA precipitate, complete medium was added and the entire suspension plated. Human umbilical vein endothelial cells, first passage, were isolated and grown as described [18].

Metabolic labeling and vWf purification

Cells were labeled for 2 to 3 days with 30 µCi/ml [³⁵S]cysteine (1300 Ci/ml, Amersham, Arlington Heights, IL/USA) and lysed as previously described [35]. vWf derivatives were immunopurified from the cell lysates and cell culture media as described [41] except that incubation with gelatin-Sepharose was omitted. Endoglycosidase H digestion of purified samples was as described [38].

Gel electrophoresis

SDS-polyacrylamide gels and SDS-agarose gels were prepared as described by Laemmli [13] and Mayadas and Wagner [18], respectively. Densitometric scanning of autoradiographs was performed, and the amount of vWf in the samples was quantitated by determining the area under the peaks.

Electron microscopy and immunocytochemistry

Cells were fixed in culture flasks with 1% glutaraldehyde in 0.1 M phosphate buffer (pH 7.4) for 1 h at 22°C, washed three times in the buffer and sent from Boston to Paris by express mail in the third wash buffer. For standard electron microscopy, the cells were post-fixed in 1% osmium tetroxide, stained en bloc with uranyl acetate, flushed off the culture flasks with propylene oxide and embedded in Epon. For immunocytochemistry, cells were embedded in glycol methacrylate after glutaraldehyde fixation [14]. The immunogold staining was performed as described [5]. Sections were incubated for 2 h in rabbit polyclonal anti-vWf at a 1:50 dilution, followed by goat anti-rabbit immunoglobulin fraction coupled to 10-nm colloidal gold, for 1 h at a 1:10 dilution. In both cases, sections were stained in uranyl acetate followed by lead nitrate and examined on a CM10 Philips electron microscope.

THIS PAGE BLANK (USPTO)

Results

Previous studies have shown that deletion of the vWf propeptide prevented vWf multimerization [32, 44] and storage [41] in transfected cells. It has also been reported that inhibition of the cleavage of the prosequence by site-specific mutagenesis of the cleavage site does not prevent vWf multimerization [33, 44]. Therefore, we were interested in finding out if blockage of cleavage effects vWf storage. For this purpose we have changed Arg 763 (the last amino acid of the prosequence) to Gly, in a variant designated pro-vWf Gly763. This mutation was shown by Verweij and colleagues [33] to completely inhibit prosequence cleavage.

The mutated cDNA was inserted into the pCMV4 expression vector, under the control of the cytomegalovirus promoter and linked at its 3' end to SV40 splice and polyadenylation signals. Multimerization and storage of the uncleavable variant were studied by expression of the pCMV4-pro-vWf Gly763 construct in heterologous cells (COS-1, AtT-20, and RIN 5F) previously shown to be suitable models for such studies [2, 41].

Inhibition of prosequence cleavage and vWf multimerization

The vWf-containing expression vectors, pCMV RlvWf (wild type; [41]) and pCMV4-pro-vWf Gly763, were transiently expressed in COS-1 cells. The cells were metabolically labeled with [³⁵S]cysteine for 2 days and the vWf-related proteins isolated from the cell lysates and culture media. Wild type pro-vWf and pro-vWf Gly763 were immunopurified using polyclonal antibodies to vWf together with a monoclonal antibody to the propolypeptide. 5% SDS-polyacrylamide gel electrophoresis analysis of the reduced samples showed that indeed the cleavage mutant molecules were not proteolytically processed (Fig. 1a; [33]). When analyzed nonreduced on 2% agarose gels (Fig. 1b), pro-vWf Gly763 multimers produced sharper bands because these were composed of precursor subunits only, while wild type multimers contained a mixture of precursor and mature subunits [16, 36]. We observed an increased efficiency of multimerization for the uncleaved mutant in that it was relatively richer in the largest multimeric species than wild type vWf. To determine this quantitatively, multimers secreted by cells from five independent transfection experiments were analyzed by scanning the autoradiographs of the agarose gels. The percent of total radioactivity in the upper fifth of the gel, where the largest multimers migrate, was determined. The mean content of these largest multimers in the five experiments was 14.4% (SEM = 1.7) for the wild type, and 25.2% (SEM = 1.3) for pro-vWf Gly763.

Inhibition of prosequence cleavage and vWf storage

In cells capable of protein storage (such as RIN 5F and AtT-20 cells), pro-vWf is diverted into the regulated pathway of secretion. We have established criteria for vWf storage in these cells [41]: the first one consists in the presence

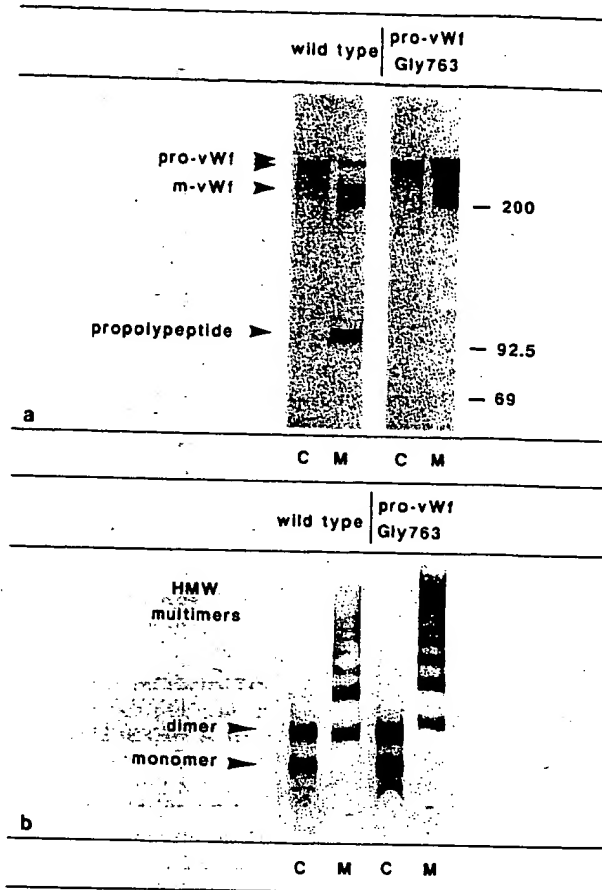


Fig. 1. Subunit and multimeric composition of pro-vWf Gly763 expressed by transfected COS-1 cells. COS-1 cells transiently expressing the wild type and pro-vWf Gly763 cDNAs were metabolically labeled with [³⁵S]cysteine for 2 days. vWf-related products were immunopurified from lysed cells (C) and culture media (M) with a polyclonal antibody to vWf, mixed (a) or not (b) with a monoclonal antibody directed against the propolypeptide. —a. Autoradiograph of a 5% polyacrylamide gel in which all the samples were simultaneously analyzed. Numbers on the right indicate the position of molecular weight standards ($\times 1000$). —b. Samples were analyzed non-reduced on a single 2% agarose gel, autoradiograph of which is shown. In (a) and (b), the lanes are as indicated on top of the gels.

of a significant intracellular pool of Endo-H-resistant vWf. This pool represents about half of total intracellular vWf in endothelial cells and in cells transfected with wild type pro-vWf [36, 41]. For its most part, the Endo-H-sensitive other half is in the endoplasmic reticulum. Only trace amounts of Endo-H-resistant material are found in cells that do not have a regulated pathway of secretion [41]. A second criterion is the presence of numerous vWf-containing granules in the transfected cells, as determined by fluorescence or electron microscopy.

Weibel-Palade bodies contain stoichiometric amounts of the mature vWf subunit and of the free propolypeptide [7], which indicates that vWf is sorted to the storage granules in

THIS PAGE BLANK (USPTO)

its pro-vWf form. To examine whether the prosequence cleavage may be linked to vWf storage in the Weibel-Palade bodies, we stably transfected the RIN 5F and the AtT-20 cells with the pCMV4-pro-vWf Gly763 plasmid. Several positive clones (3 for AtT-20 cells, and 6 for RIN 5F cells) were identified by immunofluorescence staining with an anti-vWf antiserum and expanded.

Biochemistry. The cells were metabolically labeled with [35 S]cysteine for 2 to 3 days, and immunopurified pro-vWf Gly763 was analyzed by gel electrophoresis. As expected, the secreted protein was multimeric (Fig. 2a) and similar to the COS-1 cells the mutation completely inhibited the cleavage of prosequence within the RIN 5F and the AtT-20 cells (Fig. 2b). The unexpected finding was the absence of a significant pool of Endo-H-resistant vWf in these cells

(Fig. 2b). Furthermore, high molecular weight multimers were absent from the lysed cells (Fig. 2a) while they were efficiently secreted through the constitutive pathway. These results indicated that, despite their extensive multimeric structure, the pro-vWf Gly763 molecules were not stored.

Immunofluorescence analysis. In accordance with this finding and as opposed to the wild type, Weibel-Palade body-like granules were not found in any of the isolated clones by immunofluorescence staining. We have also not observed round granules distributed throughout the cytoplasm of the transfected cells, such as were present in RIN 5F cells expressing truncated pro-vWf lacking the C-terminal 80 kDa of the mature subunit [41]. However, in some cells, speckled pattern among the strong perinuclear staining was visible which we have not observed in cells transfected with wild type pro-vWf (Fig. 3).

Electron microscopy. To obtain more detailed information on the intracellular distribution of pro-vWf Gly763 and an explanation for the perinuclear "speckled pattern" observed by immunofluorescence, we analyzed one of the transfected RIN 5F clones by ultrastructural and immunoelectron microscopy. For comparison, we chose cells expressing a truncated form of the protein lacking the 20 kDa C-terminal portion [41]. These cells were chosen because they synthesize as much vWf as the pro-vWf Gly763 clones do, whereas the amount produced by the intact wild type expressing clones is much lower (not shown). Therefore, the latter clones are not adequate for comparison. The truncated pro-vWf expressing cells display the classical morphology of the RIN 5F cell line. Their striking peculiarity is the presence of Weibel-Palade body-like structures similar to the ones of endothelial cells, showing the classical internal tubular structure ([41]; Fig. 4a). These Weibel-Palade body-like structures were found in the Golgi region and scattered in the cytoplasm.

As opposed to the pro-vWf -20 kDa expressing control cells, the pro-vWf Gly763 expressing cells did not display any Weibel-Palade body-like structures. Secretion granules were observed, as well as frequent lysosomal figures, either lamellar bodies (Figs. 5, 6b) or large membrane-bounded vacuoles containing a little electron-dense material (Fig. 6). A mean of five lamellar lysosomes per cell section was observed, whereas they were only occasionally found in control cells (approximately one per cell section).

Immunoelectron microscopy. Immunogold labeling of the control cells showed the presence of vWf in the endoplasmic reticulum, in the Golgi saccules and associated vesicles, as well as in round granules and numerous Weibel-Palade bodies (Fig. 4b). Similarly, immunolabeling for vWf in the pro-vWf Gly763 cells was present in the endoplasmic reticulum and Golgi cisternae, and in round granules (Fig. 6b). Strikingly, vWf was also frequently detected in the lysosomal structures, either lamellar bodies or large vacuoles (Figs. 5c, 6b), which was not the case for the control cells. These ultrastructural and immunocytochemical observations indicate that the pro-vWf Gly763 expressing cells display a strong and specific autophagic activity, which may be related to the heterologous mutant protein.

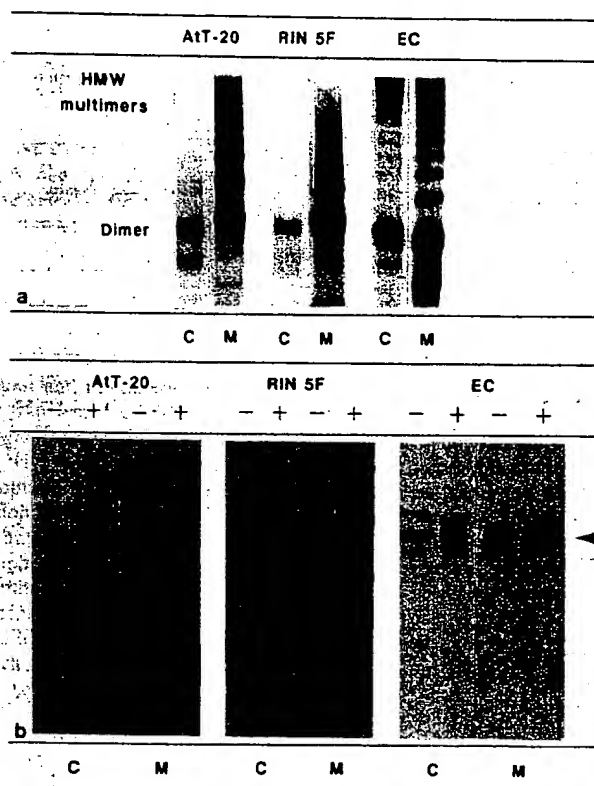


Fig. 2. Multimeric and subunit composition of pro-vWf Gly763 expressed in AtT-20 and RIN 5F cells. Cells were metabolically labeled for 2 to 3 days, and vWf-related products were immunopurified from the cell lysates (C) and the culture media (M) with an anti-vWf antiserum. —a. The samples were analyzed non-reduced on 2% agarose gels, autoradiographs of which are presented. —b. The same samples were treated with (+) or without (–) Endo-H and then analyzed reduced on 4.5% polyacrylamide gels, autoradiographs of which are shown. In both transfected cell types, less than 10% of cellular vWf was Endo-H resistant. In (a) and (b), the left and center panels show the products expressed by AtT-20 and RIN 5F cells, respectively. The right panel shows wild type vWf from endothelial cells (EC) for comparison.

THIS PAGE BLANK (USPTO)

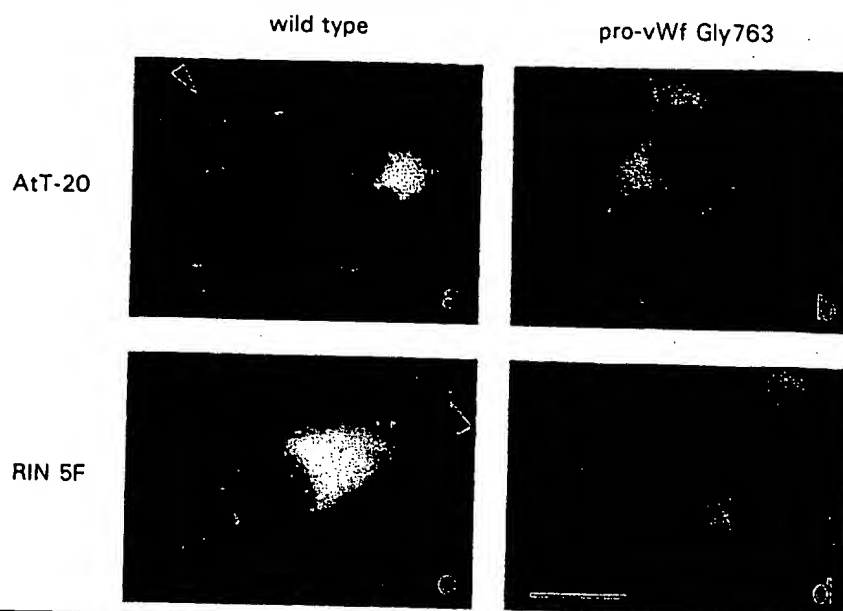


Fig. 3. Intracellular distribution of pro-vWf Gly763 in AtT-20 and RIN 5F cells. Immunofluorescence staining with anti-vWf antiserum of permeabilized wild type pro-vWf expressing AtT-20 cells (a) and RIN 5F cells (c); pro-vWf Gly763 expressing AtT-20

cells (b) and RIN 5F cells (d). Arrowheads indicate Weibel-Palade body-like structures in the wild type transfected cells. Pro-vWf Gly763 did not form distinct granules in either cell type. — Bar 10 μ m.

Discussion

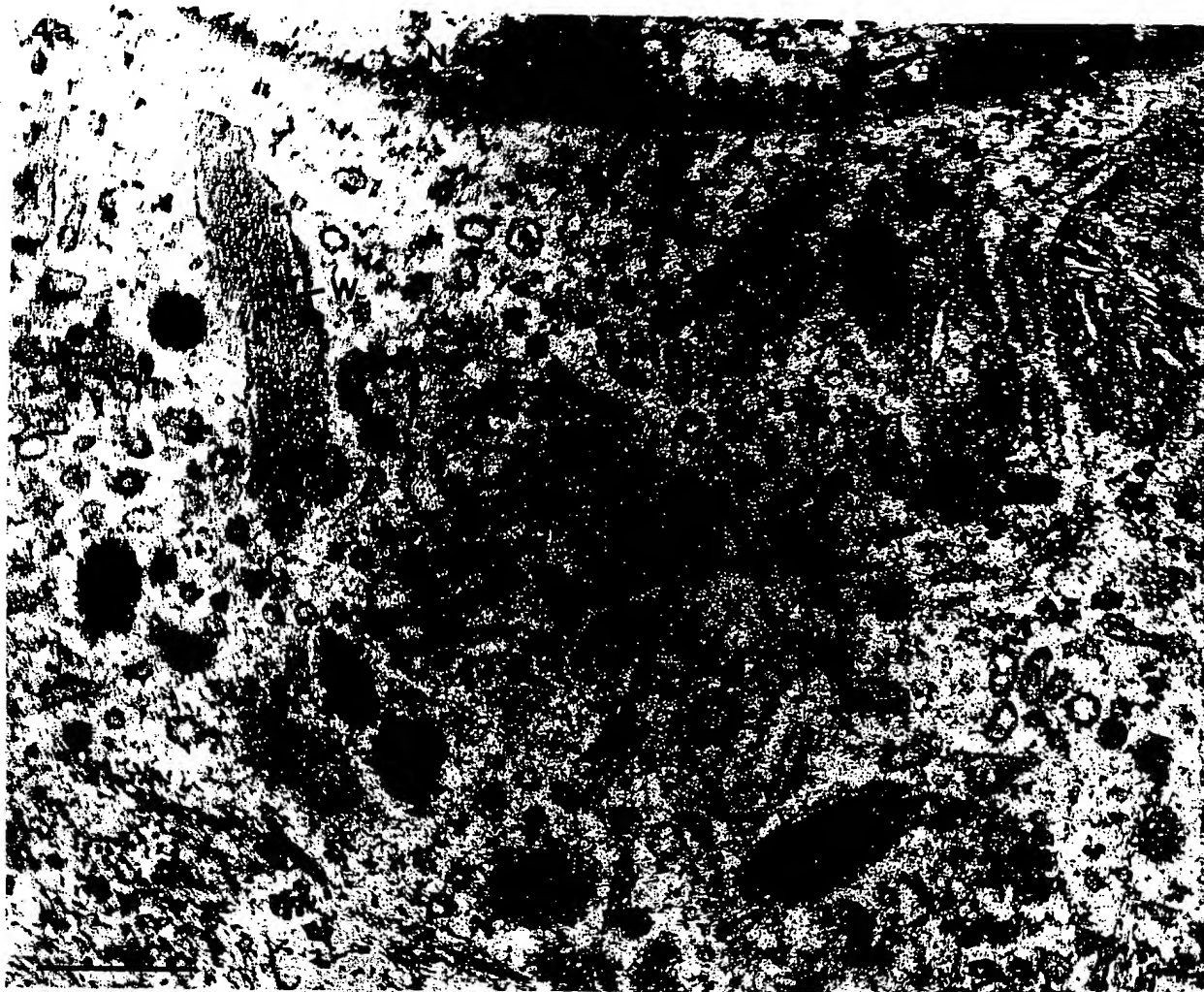
The large vWf propolypeptide has a role in both vWf processing and intracellular transport. It is required for multimerization of pro-vWf dimers [18, 32, 44] and for storage of the multimeric protein in specific storage granules [41]. To investigate the importance of proteolytic processing on vWf storage, we have completely inhibited prosequence cleavage by replacing Arg adjacent to the cleavage site with Gly, in a mutant designated pro-vWf Gly763 [33]. Similar studies on vWf mutants have proved the validity of expression systems using COS-1 cells [2] AtT-20 and RIN 5F cells [41] as recipient cells for vWf cDNA derivatives. Therefore, we chose these three cell lines to analyze the pattern of multimerization and the extent of storage of this vWf variant.

Multimerization is thought to be initiated by non-covalent interactions between propolypeptides of separate pro-vWf dimers or oligomers (reviewed in [19]). It is believed that these interactions promote the correct positioning of amino-terminal parts of mature vWf such that disulfide bonding can occur. This hypothesis is supported by the non-covalent dimeric nature of the free propolypeptide in plasma [39]. Although multimerization of the pro-vWf Gly763 mutant was more efficient than wild type, it did not proceed to completion, i. e., the typical array of small and large vWf multimers was secreted by the transfected cells (Figs. 1b, 2a). This indicates that the cleavage of the prosequence is not the major reason for arrest of multimerization and that other cellular factors, like local pro-vWf di-

mer concentration or time taken by the protein to reach the extracellular environment are regulating the extent of multimer formation.

Despite the formation of large disulfide-bonded multimers up to 20 million daltons in size (Fig. 2a), we did not detect storage of pro-vWf Gly763 in neither AtT-20 nor RIN 5F cells, according to the previously established criteria [41]. There was no significant Endo-H-resistant vWf pool in these cells (Fig. 2b) nor did we observe storage granules by fluorescence microscopy (Fig. 3). In both cell types, wild type or truncated vWf species form their own storage granules (Fig. 3) that are distinct from the endogenous hormone-containing granules [41]. While the wild type vWf-containing granules have a rod shape similar to Weibel-Palade bodies, granules formed by a 80 kDa C-terminal deletion mutant are round [41]. Therefore, we feel we would detect vWf-containing granules even of unusual shape if they were present in significant numbers. In addition, we did not observe accumulation of vWf staining at the tips of AtT-20 cells which would have been indicative of cotargeting of pro-vWf Gly763 with adrenocorticotrophic hormone [4, 12, 26, 29]. By immunoelectron microscopy, we have detected pro-vWf Gly763 in a few round granules and in numerous lysosomal structures. These lysosomal structures may be the reason for the speckled pattern seen in the perinuclear region by fluorescence staining. The control cell contained about 5 times fewer lysosomes in which label for vWf was not detected. Weibel-Palade body-like structures were found only in the control cells. It appears therefore that although rare granules form in the pro-vWf

THIS PAGE BLANK (USPTO)



Gly763 cells, these may be short-lived and the majority of the vWf produced by these cells is either secreted constitutively or degraded in lysosomes.

Another example of an uncleavable prohormone which was not efficiently stored was described by Orci and colleagues [23]: they showed that growing pancreatic B cells in the presence of Arg and Lys analogs rendered the synthesized proinsulin resistant to prohormone processing. The uncleaved prohormone remained in immature secretory granules and was degraded more rapidly than the native insulin [9]. A similar phenomenon could prevent storage of the uncleavable pro-vWf Gly763 and direct it to lysosomal structures.

The lack of efficient storage of the pro-vWf Gly763 mutant may be explained in at least two ways. First, removal of the propolypeptide could be necessary for storage. Pro-vWf [7, 39] similar to prohormones [17, 24] appears to be targeted to the storage granules in the precursor form since in the granules the mature protein and the propeptide are

found in stoichiometric amounts. The proteolytic processing of the stored proteins is very efficient. For example, no pro-vWf subunits are detected in Weibel-Palade bodies releasate whereas a large portion of constitutively secreted protein is not cleaved [28]. Perhaps the formation of the tubular structures containing vWf, as found in Weibel-Palade bodies [42], requires the prosequence cleavage, and it is only these structures that can be successfully stored. This would be analogous to collagen type I fibril assembly which is initiated by removal of terminal propeptides by cleavage of the procollagen molecules with specific proteinases [10]. Nevertheless, it would be surprising that the quaternary structure of the protein would be so important for storage. As we have already mentioned, deletion of the C-terminal 80 kDa from pro-vWf did not interfere with its storage in granules whose morphology did not resemble that of Weibel-Palade bodies [41]. In addition, chimeric proteins composed of a usually stored and a non-stored component were successfully directed to the regulated pathway of secretion

THIS PAGE BLANK (USPTO)

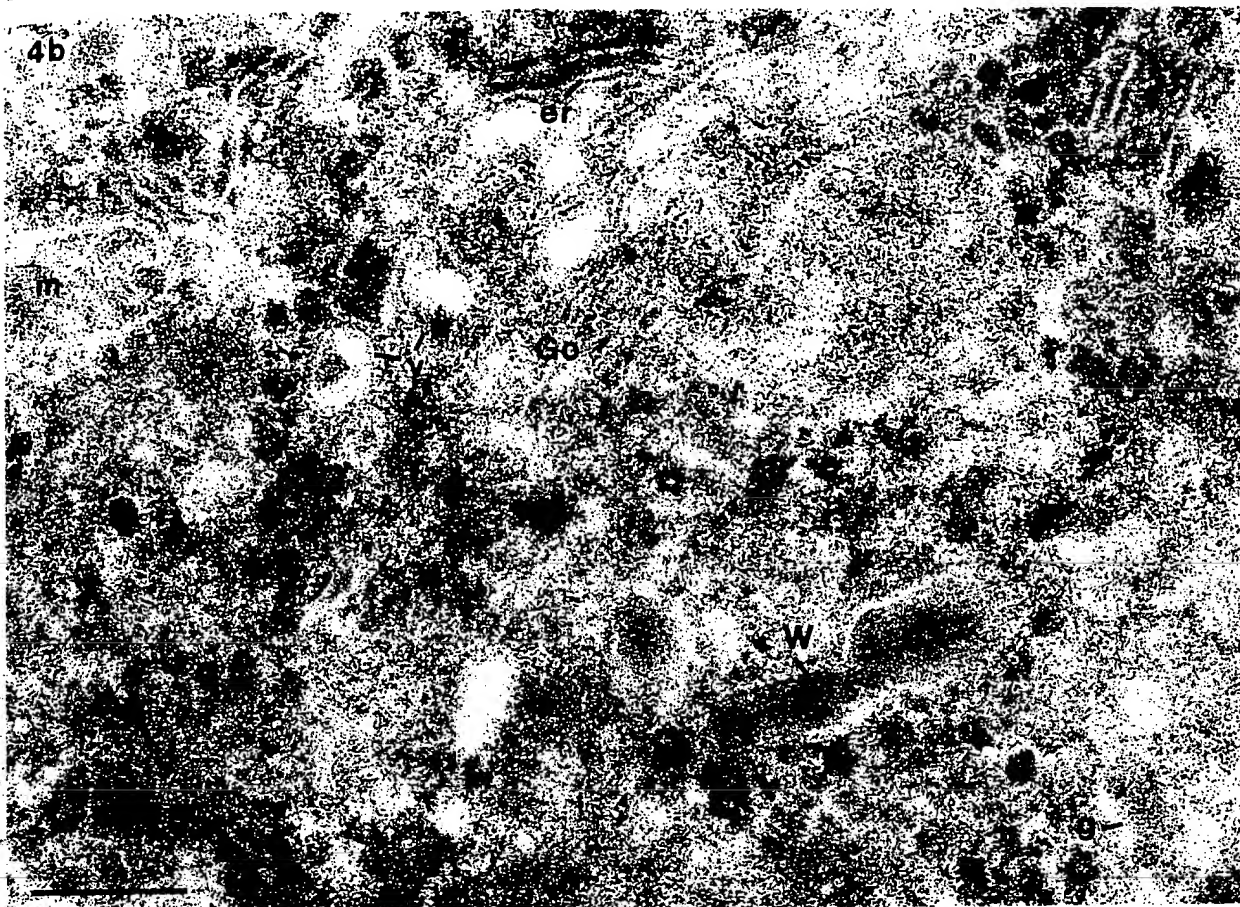


Fig. 4. Electron micrographs of a RIN 5F cell expressing the 20-kDa pro-vWf species: —a. Standard electron microscopy analysis. Typical Weibel-Palade bodies (W) with aligned tubular structures are present in the centriolar region (ce), together with other granules (g) near the Golgi cisternae (Go). —er Endoplasmic reticulum.

—m Mitochondrion. —N Nucleus. —b. Immunogold labeling for vWf. Gold particles are present in the endoplasmic reticulum (er), Golgi cisternae (Go) and associated vesicles (v), in some round cytoplasmic granules (g), and in typically striated Weibel-Palade bodies (W). —Bars 0.4 μ m.

[22, 27]. These proteins would be expected to exhibit anomalous quaternary structure.

A second possibility is that the cleavage site could be part of a targeting sequence that would not be recognized anymore or not recognized efficiently upon mutation. Although there are no obvious sequence homologies among stored proteins, they are likely recognized by a common mechanism that is preserved across species and tissues [11]. The majority of the stored proteins, including vWf, have two characteristics in common: they form aggregates, and they are proteolytically processed at the C-terminal side of basic residues. This processing is performed by a class of proteins: the prosequence cleavage enzymes. A mammalian member (furin/PACE) of the family of proteases homologous to the yeast gene product of KEX2 [21] was recently cloned [30]. The functionality of the enzyme was tested in COS-1 cells by cotransfection of its cDNA with the pro-vWf cDNA [31, 45] or the pro-vWf Gly763 cDNA [31]. The recombinant protease was able to correctly cleave the

wild type pro-vWf, whereas pro-vWf Gly763 was resistant to this proteolytic activity. Two other enzymes belonging to the same family of proteases, PC2 and PC3, were unable to cleave pro-vWf [25]. Their known substrates do not contain an Arg at the -4 position from the cleavage site, as opposed to all the potential substrates for furin/PACE, including vWf. This suggests that the specific processing enzyme for vWf is furin/PACE. Furin/PACE has a potential transmembrane domain which makes it a good candidate for a targeting receptor. Upon cleavage of the prosequence which occurs in the immature secretory granules [24], such a processing enzyme/targeting receptor might dissociate from the protein products, allowing it to recycle, whereas the cleaved hormone would go to the storage granule.

In conclusion, the lack of storage of pro-vWf Gly763 demonstrates that the capacity of a protein to form large aggregates is not sufficient to elicit its storage in cells with a regulated pathway of secretion. Although aggregation of vWf may play an important role in the formation of Wei-

THIS PAGE BLANK (USPTO)

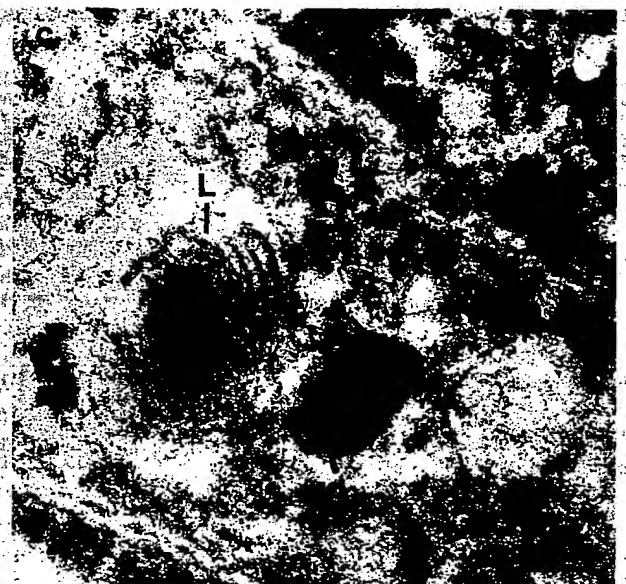
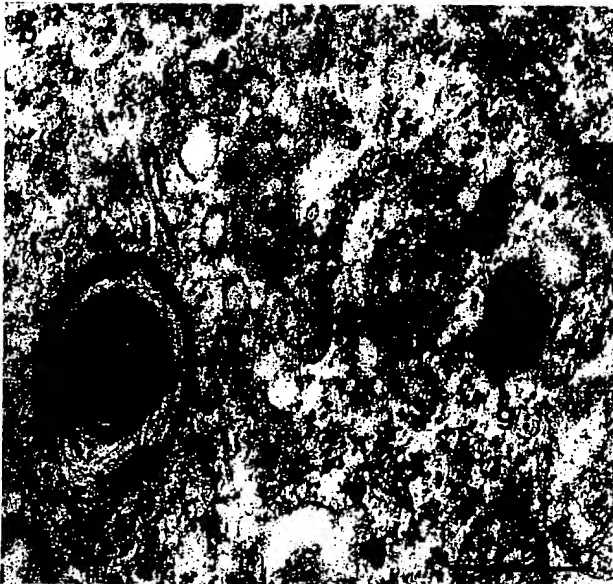
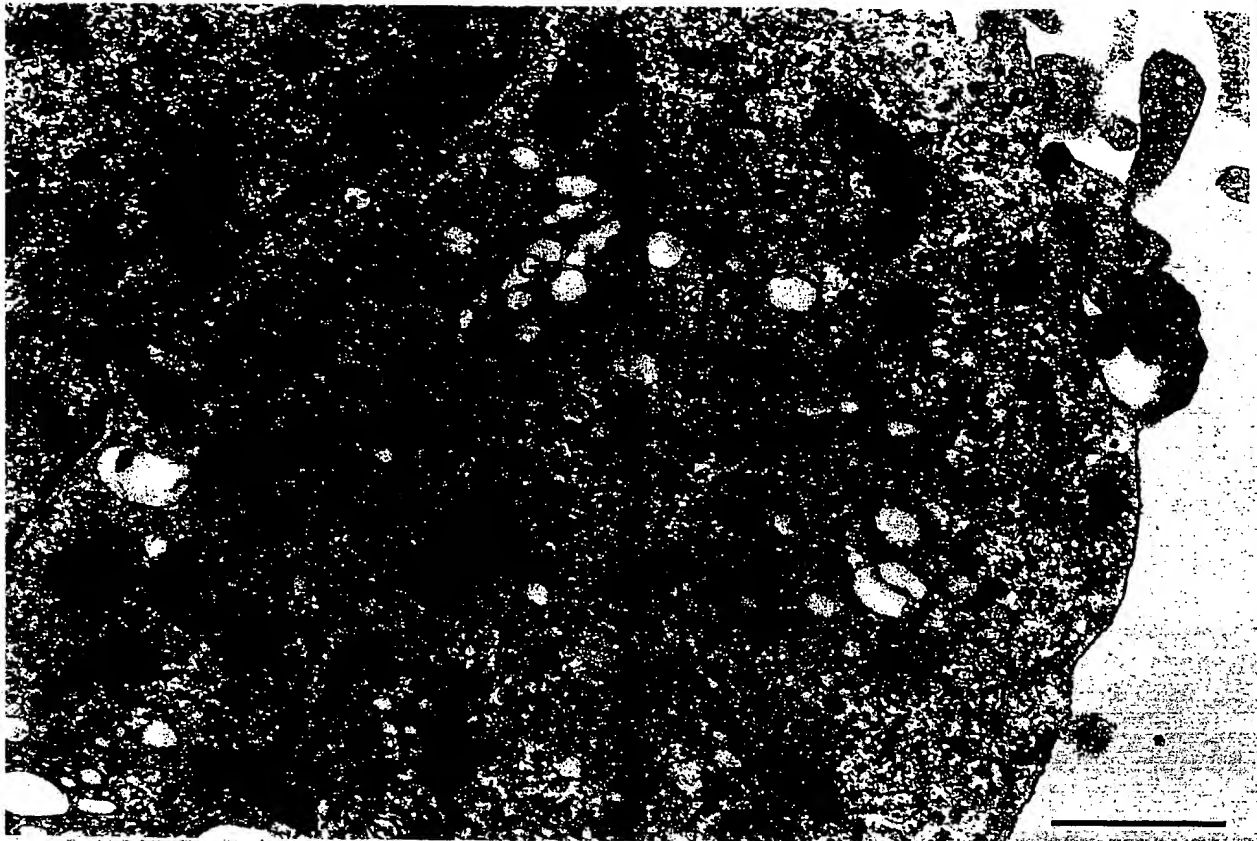
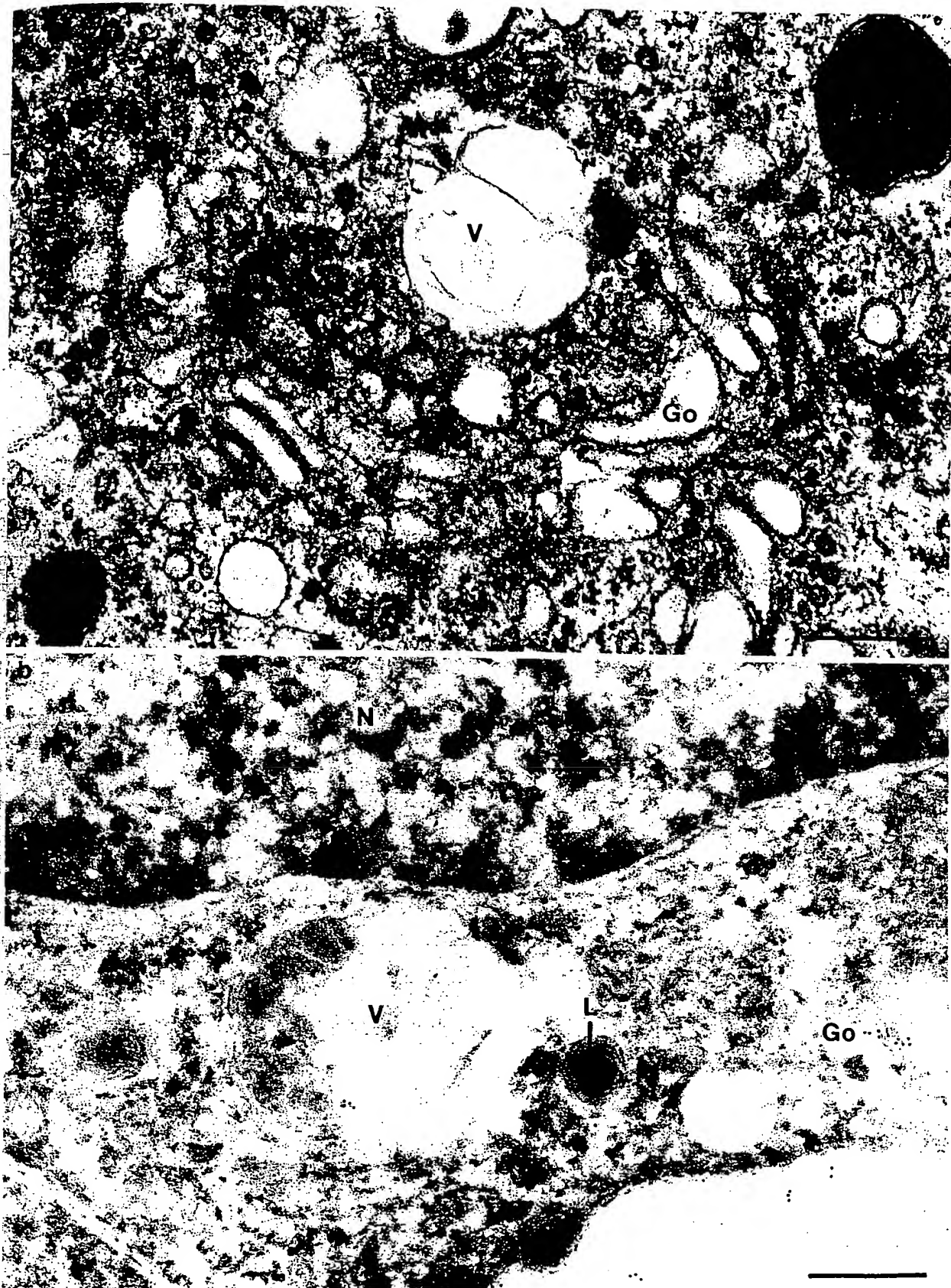


Fig. 5. Electron micrographs of a pro-vWf Gly763 expressing RIN 5F cell. —a. Global view of the cell by standard electron microscopy. The cell displays the usual aspect of RIN 5F cells with prominent Golgi complexes (Go) located in the centriolar region (ce). Some secretion granules and, more specifically to these transfected cells, numerous lysosomes (L) are present. —m Mitochondria. —N Nucleus. —b. Enlargement of a lysosomal structure (L) with concentric layers, and secretion granules (g). —c. Immunogold staining for vWf. Gold particles are present in some

round secretion granules (g) and lysosomes (L). —Bars 1 μ m (a), 0.4 μ m (b), 0.2 μ m (c).

Fig. 6. High magnification electron micrographs of a pro-vWf Gly763 expressing RIN 5F cell. —a. Enlargement of the Golgi zones (Go) with some large clear vacuoles (V). —b. Immunogold staining for vWf. These vacuoles (V) occasionally label for vWf, and the lysosomal structures (L) label as well. —Go Golgi complex. —N Nucleus. —Bars 0.4 μ m.

THIS PAGE BLANK (USPTO)



THIS PAGE BLANK (USPTO)

bel-Palade bodies, additional factors such as presence of the prosequence and of an intact prosequence cleavage site are required. In future studies, the role of the prosequence cleavage enzymes as putative targeting receptors should be evaluated.

Acknowledgements. The authors thank Margaret Jacobs for oligonucleotide synthesis and Drs. Tanya N. Mayadas and Joost Koedam for helpful discussions. — This work was supported by a National Institute of Health grant RO1 HL41002 and PO1 HL42443, and funded by the Association pour la Recherche sur le Cancer for the electron microscopy part.

References

- [1] Bonfanti, R., B. C. Furie, B. Furie, D. D. Wagner: PADGEM (GMP140) is a component of Weibel-Palade bodies of human endothelial cells. *Blood* 73, 1109–1112 (1989).
- [2] Bonthron, D. T., R. I. Handin, R. J. Kaufman, L. C. Wasley, E. C. Orr, L. M. Mitsock, B. Ewenstein, J. Loscalzo, D. Ginsburg, S. H. Orkin: Structure of pre-pro-von Willebrand factor and its expression in heterologous cells. *Nature* 324, 270–273 (1986).
- [3] Bonthron, D., E. C. Orr, L. M. Mitsock, D. Ginsburg, R. I. Handin, S. H. Orkin: Nucleotide sequence of pre-pro-von Willebrand factor cDNA. *Nucl. Acids Res.* 14, 7125–7127 (1986).
- [4] Burgess, T. L., C. S. Craik, L. Matsuchi, R. B. Kelly: In vitro mutagenesis of trypsinogen: role of the amino terminus in intracellular protein targeting to secretory granules. *J. Cell Biol.* 105, 659–668 (1987).
- [5] Cramer, E. M., D. Meyer, R. Le Menn, J. Breton-Gorius: Eccentric localization of von Willebrand factor within a tubular structure of platelet alpha-granule resembling that of Weibel-Palade bodies. *Blood* 66, 710–713 (1985).
- [6] Disdier, M., J. H. Morrissey, R. D. Fugate, D. F. Bainton, R. P. McEver: Cytoplasmic domain of P-selectin (CD 62) contains the signal for sorting into the regulated secretory pathway. *Mol. Biol. Cell* 3, 309–322 (1992).
- [7] Ewenstein, B. M., M. J. Warhol, R. I. Handin, J. S. Pober: Composition of the von Willebrand factor storage organelle (Weibel-Palade body) isolated from cultured human umbilical vein endothelial cells. *J. Cell Biol.* 104, 1423–1433 (1987).
- [8] Graham, F. L., A. J. van der Eb: A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52, 456–467 (1973).
- [9] Halban, P. A., M. Amherdt, L. Orci, A. E. Renold: Proinsulin modified by analogs of arginine and lysine is degraded rapidly in pancreatic B cells. *Biochem. J.* 219, 91–97 (1984).
- [10] Kadler, K. E., Y. Hojima, D. J. Prockop: Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. *J. Biol. Chem.* 260, 15696–15701 (1987).
- [11] Kelly, R. B.: Pathway of protein secretion in eukaryotes. *Science* 230, 25–32 (1985).
- [12] Koedam, J. A., E. M. Cramer, E. Briend, B. Furie, B. C. Furie, D. D. Wagner: P-selectin, a granule membrane protein of platelets and endothelial cells, follows the regulated secretory pathway in AtT-20 cells. *J. Cell Biol.* 116, 617–625 (1992).
- [13] Laemmli, U. K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685 (1970).
- [14] Leduc, E. M., W. Bernhard: Recent modification of the glycol-methacrylate embedding procedure. *J. Ultrastruct. Res.* 19, 196–199 (1967).
- [15] Lynch, D. C., R. Williams, T. S. Zimmerman, E. P. Kirby, D. M. Livingston: Biosynthesis of the subunits of factor VIIIr by bovine aortic endothelial cells. *Proc. Natl. Acad. Sci. USA* 80, 2738–2742 (1983).
- [16] Lynch, D. C., T. S. Zimmerman, E. H. Ling, P. J. Browning: An explanation for minor multimer species in endothelial cell-synthesized von Willebrand factor. *J. Clin. Invest.* 77, 2048–2053 (1986).
- [17] Mains, R. E., B. A. Eipper: Peptide products derived from pro- ACTH/endorphin by mouse pituitary tumor cells. *J. Cell Biol.* 89, 21–28 (1981).
- [18] Mayadas, T. N., D. D. Wagner: In vitro multimerization of von Willebrand factor is triggered by low pH. Importance of the propolypeptide and free sulfhydryls. *J. Biol. Chem.* 264, 13497–13503 (1989).
- [19] Mayadas, T. N., D. D. Wagner: von Willebrand factor biosynthesis and processing. *Ann. NY Acad. Sci.* 614, 153–166 (1991).
- [20] McEver, R. P., J. H. Beckstead, K. L. Moore, L. Marshall-Carlson, D. F. Bainton: GMP-140, a platelet alpha-granule membrane protein, is also synthesized by vascular endothelial cells and is localized in Weibel-Palade bodies. *J. Clin. Invest.* 84, 92–99 (1989).
- [21] Mizuno, K., T. Nakamura, T. Ohshima, S. Tanaka, H. Matsuo: Yeast KEX2 gene encodes an endopeptidase homologous to subtilisin-like serine proteases. *Biochem. Biophys. Res. Commun.* 156, 246–254 (1988).
- [22] Moore, H.-P. H., R. B. Kelly: Re-routing of a secretory protein by fusion with human growth hormone sequences. *Nature* 321, 443–446 (1986).
- [23] Orci, L., P. Halban, M. Amherdt, M. Ravazzola, J.-D. Vassalli, A. Perrelet: Nonconverted, amino acid analog-modified proinsulin stays in a Golgi-derived clathrin-coated membrane compartment. *J. Cell Biol.* 99, 2187–2192 (1984).
- [24] Orci, L., M. Ravazzola, M. Amherdt, O. Madsen, J.-D. Vassalli, A. Perrelet: Direct identification of prohormone conversion site in insulin-secreting cells. *Cell* 42, 671–681 (1985).
- [25] Rehemtulla, A., R. J. Kaufman: Preferred sequence requirements for cleavage of pro-von Willebrand factor by propeptide-processing enzymes. *Blood* 79, 2349–2355 (1992).
- [26] Rivas, R. J., H.-P. H. Moore: Spatial segregation of the regulated and constitutive pathways. *J. Cell Biol.* 109, 51–60 (1989).
- [27] Seethaler, G., M. Chaminade, R. Vlasak, M. Ericsson, G. Griffiths, O. Toffoletto, J. Rossier, H. G. Stunnenberg, G. Kreil: Targeting of frog prodermorphin to regulated secretory pathway by fusion to proenkephalin. *J. Cell Biol.* 114, 1125–1133 (1991).
- [28] Sporn, L. A., V. J. Marder, D. D. Wagner: Inducible secretion of large, biologically potent von Willebrand factor multimers. *Cell* 46, 185–190 (1986).
- [29] Tooze, J., B. Burke: Accumulation of adrenocorticotropin secretory granules in the midbody of telophase AtT-20 cells: evidence that secretory granules move anterogradely along microtubules. *J. Cell Biol.* 104, 1047–1057 (1987).
- [30] Van den Ouweland, A. M. W., H. L. P. van Duijnhoven, G. D. Keizer, L. C. J. Dorssers, W. J. M. Van de Ven: Structural homology between the human fur gene product and the subtilisin-like protease encoded by yeast KEX2. *Nucl. Acids Res.* 18, 664 (1990).
- [31] Van de Ven, W. J. M., J. Voorberg, R. Fontijn, H. Pannekoek, A. M. W. van den Ouweland, H. L. P. van Duijnhoven, A. J. M. Roebroek, R. J. Siezen: Furin is a subtilisin-like proprotein processing enzyme in higher eukaryotes. *Mol. Biol. Rep.* 14, 265–275 (1990).
- [32] Verweij, C. L., M. Hart, H. Pannekoek: Expression of variant von Willebrand factor (vWf) cDNA in heterologous cells: Re-

THIS PAGE BLANK (USPTO)

- quirement of the pro-polypeptide in vWf multimer formation. *EMBO J.* 6, 2885-2890 (1987).
- [33] Verweij, C. L., M. Hart, H. Pannekoek: Proteolytic cleavage of the precursor of von Willebrand factor is not essential for multimer formation. *J. Biol. Chem.* 263, 7921-7924 (1988).
- [34] Wagner, D. D.: Cell biology of von Willebrand factor. *Annu. Rev. Cell Biol.* 6, 217-246 (1990).
- [35] Wagner, D. D., V. J. Marder: Biosynthesis of von Willebrand protein by human endothelial cells. Identification of a large precursor polypeptide chain. *J. Biol. Chem.* 258, 2065-2067 (1983).
- [36] Wagner, D. D., V. J. Marder: Biosynthesis of von Willebrand protein by human endothelial cells: Processing steps and their intracellular localization. *J. Cell Biol.* 99, 2123-2130 (1984).
- [37] Wagner, D. D., J. B. Olmsted, V. J. Marder: Immunolocalization of von Willebrand protein in Weibel-Palade bodies of human endothelial cells. *J. Cell Biol.* 95, 355-360 (1982).
- [38] Wagner, D. D., T. Mayadas, V. J. Marder: Initial glycosylation and acidic pH in the Golgi apparatus are required for multimerization of von Willebrand factor. *J. Cell Biol.* 102, 1320-1324 (1986).
- [39] Wagner, D. D., P. J. Fay, L. A. Sporn, S. Sinha, S. O. Lawrence, V. J. Marder: Divergent fates of von Willebrand factor and its propeptide (von Willebrand antigen II) after secretion from endothelial cells. *Proc. Natl. Acad. Sci. USA* 84, 1955-1959 (1987).
- [40] Wagner, D. D., S. O. Lawrence, B. M. Ohlsson-Wilhelm, P. J. Fay, V. J. Marder: Topology and order of formation of interchain disulfide bonds in von Willebrand factor. *Blood* 69, 27-32 (1987).
- [41] Wagner, D. D., S. Saffaripour, R. Bonfanti, J. E. Sadler, E. M. Cramer, B. Chapman, T. N. Mayadas: Induction of specific storage organelles by von Willebrand factor propeptide. *Cell* 64, 403-413 (1991).
- [42] Weibel, E. R., G. E. Palade: New cytoplasmic components in arterial endothelia. *J. Cell Biol.* 23, 101-112 (1964).
- [43] Wigler, M., A. Pellicer, S. Silverstein, R. Axel: Biochemical transfer of single-copy eucaryotic genes using total cellular DNA as donor. *Cell* 14, 725-731 (1978).
- [44] Wise, R. J., D. D. Pittman, R. J. Handin, R. J. Kaufman, S. H. Orkin: The propeptide of von Willebrand factor independently mediates the assembly of von Willebrand multimers. *Cell* 52, 229-236 (1988).
- [45] Wise, R. J., P. J. Barr, P. A. Wong, M. C. Kiefer, A. J. Brake, R. J. Kaufman: Expression of a human proprotein processing enzyme: correct cleavage of the von Willebrand factor precursor at a paired basic amino acid site. *Proc. Natl. Acad. Sci. USA* 87, 9378-9382 (1990).

THIS PAGE BLANK (USPTO)

Mutations within the propeptide, the primary cleavage site or the catalytic site, or deletion of C-terminal sequences, prevents secretion of proPC2 from transfected COS-7 cells

Neil A. TAYLOR*, Kathleen I. J. SHENNAN*, Daniel F. CUTLER† and Kevin DOCHERTY*†

*Department of Molecular and Cell Biology, University of Aberdeen, Institute of Medical Sciences, Aberdeen AB25 2ZD, U.K., and †MRC Laboratory of Molecular Cell Biology, Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, U.K.

PC2 is a neuroendocrine endoprotease involved in the processing of prohormones and proneuropeptides. PC2 is synthesized as a proenzyme which undergoes proteolytic maturation within the cellular secretory apparatus. Cleavage occurs at specific sites to remove the N-terminal propeptide. The aim of the present study was to investigate structural requirements for the transfer of proPC2 through the secretory pathway. A series of mutant proPC2 constructs were transfected into COS-7 cells and the fate of the expressed proteins followed by pulse-chase analysis and immunocytochemistry. Human PC2 was secreted relatively slowly, and appeared in the medium primarily as proPC2 (75 kDa), together with much lower amounts of a processed intermediate (71 kDa) and mature PC2 (68 kDa). Mutations

within the primary processing site or the catalytic triad caused the protein to accumulate intracellularly, whereas deletion of part of the propeptide, the P-domain or the C-terminal regions also prevented secretion. Immunocytochemistry showed that wild-type hPC2 was localized mainly in the Golgi, whereas two representative mutants showed a distribution typical of proteins resident in the endoplasmic reticulum. The results suggest that proenzyme processing is not essential for secretion of PC2, but peptides containing mutations that affect the ability of the propeptide (and cleavage sites) to fold within the catalytic pocket are not transferred beyond the early stages of the secretory pathway. C-terminal sequences may be involved in stabilizing such conformations.

INTRODUCTION

PC2 [1,2] is a member of the eukaryotic subtilisin-like endopeptidase family. Other members of this family include the yeast enzymes Kex2 [3,4] and Krp [5], as well as furin [6], PACE4 [7], PC3 (also known as PC1) [8,9], PC4 [10], PC5/PC6 [11] and PC7 (also known as LPC or PC8) [12–14]. These proteases are involved in the proteolytic cleavage of precursor proteins at specific sites to generate mature proteins during transfer through the secretory pathway.

All members of the eukaryotic subtilisin-like endopeptidase family share certain structural features. They contain an N-terminal signal peptide, a propeptide, a subtilisin-like catalytic domain, a large conserved domain named the P-domain or HomoB domain [15], and a C-terminal tail which varies in length and may contain cysteine- or threonine-rich regions (furin, PACE4, PC5), an amphipathic helix (PC2, PC3) or a transmembrane domain (Kex2, Krp, furin, and an alternatively spliced form of PC5/6 called PC6B). The role of the P-domain is unknown, but it has been shown to be essential for production of active protein, since small deletions in this region completely abolish Kex2 and furin enzyme activity [16,17].

During transit through the secretory pathway, the subtilisin-like endoproteases are converted into the mature active form by proteolytic removal of the N-terminal propeptide. In the case of human PC2 (hPC2), the propeptide contains two processing sites, cleavage at either of which can occur independently. Processing at the sequence Arg-Lys-Lys-Arg⁵⁴ is autocatalytic [18,19], with a pH optimum and calcium requirement similar to those for substrate processing. Processing at the sequence Lys-Arg-Arg-Arg⁵⁶ can be catalysed by another enzyme, possibly furin or PACE4. Proteolytic cleavage of the propeptide of PC2

is a relatively slow process (half-life approx. 8 h), and the conditions under which maturation occurs *in vitro* suggest that *in vivo* the propeptide is removed in a late secretory compartment [19].

The role of the N-terminal propeptide in PC2 is unknown. Possible functions include: (i) acting as an intramolecular chaperone in a manner similar to the propeptide of subtilisin BPN' [20]; (ii) acting as an inhibitor that prevents premature activation of the endopeptidase; and (iii) involvement in sorting to the regulated secretory pathway. We have previously shown that proPC2 undergoes calcium- and pH-dependent aggregation and membrane association, and that a peptide corresponding to amino acids 57–85 of the propeptide can compete with proPC2 for membrane association [21]. The aim of the present study was to investigate the role of the propeptide of proPC2 in the ability of the latter to traverse the secretory pathway.

MATERIALS AND METHODS

DNA manipulations

DNA mutagenesis was carried out on an hPC2 cDNA (kindly supplied by Dr. D. F. Steiner, Howard Hughes Medical Centre, University of Chicago, IL, U.S.A.). (Δ587–613)hPC2, V54R,L56R-hPC2, (Δ81–84)hPC2, D142N-hPC2 and KDE1-hPC2 were all generated as described previously [19]. (Δ5–49)hPC2 was generated by digesting the wild-type cDNA with *Age*I and *Eco*NI, making the DNA ends flush with T4 DNA polymerase and re-ligating. (Δ54–77)hPC2 was generated by loopout mutagenesis using an oligonucleotide with the sequence 5' GTAACCTGCCTTTTTCGGTCAAATCCCTTTG-CAAGGCCATTGTGATAAAA 3'. 414-hPC2, 569-hPC2 and 573-hPC2 were generated by amplifying hPC2 cDNA using a 3'

Abbreviations used: DMEM, Dulbecco's modified Eagle's medium; Endo H, endoglycosidase H; ER, endoplasmic reticulum; hPC2, human PC2; P,N-Gase F, P,N-glycanase F.

† To whom correspondence should be addressed.

THIS PAGE BLANK (USPTO)

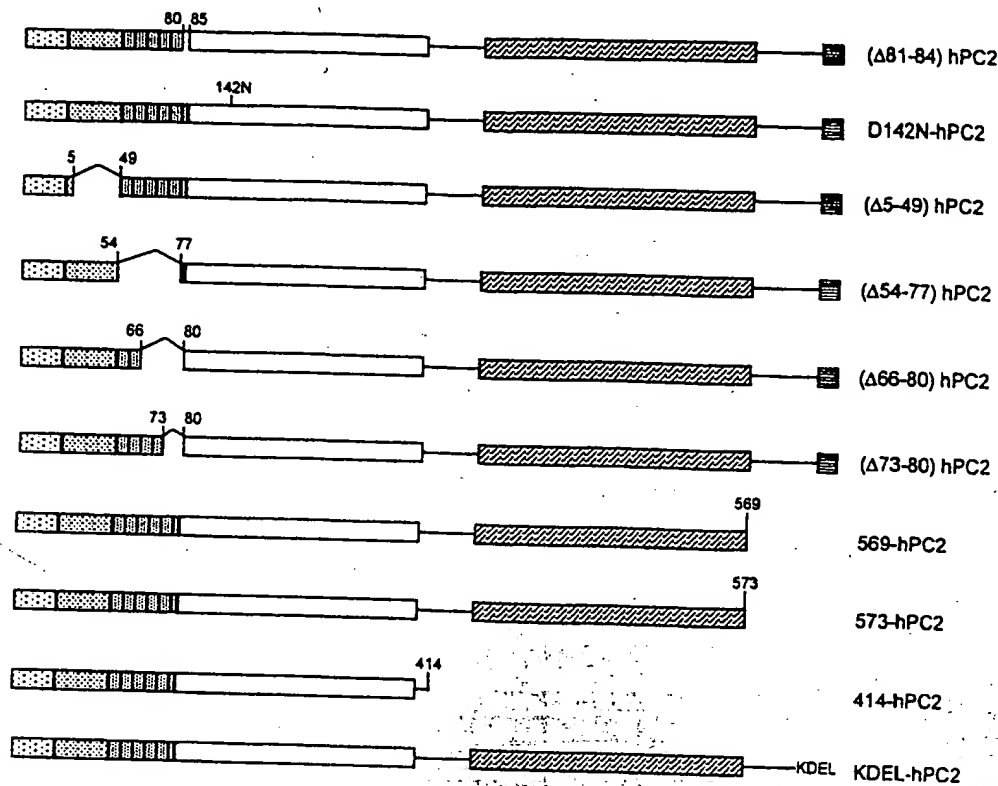


Figure 1 Structures of normal and mutant hPC2 peptides

Diagram showing the structures of hPC2 and the mutant constructs used in this study. The domains of PC2 are indicated, along with the catalytically important residues and the primary and secondary cleavage sites. Constructs were generated as described in the Materials and methods section.

oligonucleotide containing a stop codon after the codon for the residue indicated and a 5' oligonucleotide encoding the N-terminus of the preproprotein. The PCR products were cloned and the region of interest sequenced and subcloned back into the wild-type cDNA. (Δ66-80)hPC2 and (Δ73-80)hPC2 were generated using the Excite® kit (Stratagene).

Cell culture and lipofection

COS-7 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) foetal calf serum in an atmosphere of 5% CO₂. Cells at about 40% confluence in six-well plates were transfected by mixing 2 μg of DNA and 24 μl of a 1 nM lipid suspension containing a 2:1 (molar ratio) mixture of dioleoyl-L-α-phosphatidylethanolamine (Sigma, Poole, Dorset, U.K.) and dimethyldioctadecylammonium bromide (Fluka) in 0.5 ml of serum-free DMEM. The lipid-DNA complexes were allowed to form for 20 min and then added to the washed cells for 4 h before serum was replaced [22].

Pulse-chase and immunoprecipitation

At 48 h post-transfection the cells were incubated in methionine- and serum-free DMEM (Sigma) for 30 min. The medium was then removed and the cells pulse-labelled in 0.5 ml of the same medium supplemented with 170 μCi of [³⁵S]methionine (> 1000 Ci/mmol; Amersham) per well. After the 30 min pulse, the labelling medium was removed and replaced by 1 ml of complete medium containing 30 mg/l unlabelled methionine. Medium was removed at the appropriate time and the cells harvested in NDET [1% (v/v) Nonidet P40, 0.4% (w/v)

deoxycholate, 66 mM EDTA, 10 mM Tris/HCl, pH 7.4). Cell extracts were incubated on ice for 5 min, and then the nuclei were centrifuged (13000 g, 1 min) and the post-nuclear supernatant saved. hPC2 immunoreactivity was immunoprecipitated in NDET/0.3% SDS using anti-PC2 antiserum raised to amino acids-Met²³³-Asn⁶¹³ of hPC2 expressed in *Escherichia coli* using the pEt12 vector.

Endoglycosidase digestion

Immunoprecipitated material was denatured by heating to 100 °C in 0.5% SDS and 1% β-mercaptoethanol. Half of the material was incubated at 37 °C for 2 h with 100 units of endoglycosidase H (Endo H) in 50 mM sodium citrate, pH 5.5. The other half was incubated at 37 °C for 2 h with 100 units of P,N-glycanase F (P,N,Gase F) in 50 mM sodium phosphate, pH 7.5, and 1% Nonidet P40. P,N,Gase F and Endo H were supplied by New England Biolabs (Hitchin, Herts., U.K.).

Immunocytochemistry

Immunocytochemistry was carried out as detailed in Cramer and Cutler [23] using the rabbit PEP4 anti-PC2 antibody (kindly supplied by Dr. D. F. Steiner, Howard Hughes Medical Centre, University of Chicago, IL, U.S.A.).

RESULTS

The aim of this study was to investigate the structural requirements for the transfer of proPC2 through the secretory apparatus. The approach adopted was to use site-directed mutagenesis to

THIS PAGE BLANK (USPTO)

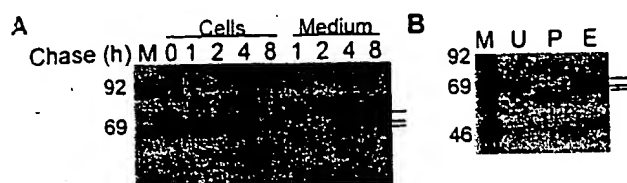


Figure 2 Expression of hPC2 in COS-7 cells

(A) Pulse-chase analysis of hPC2 immunoreactivity found in cell extracts and medium from cells transfected with pCR-PC2. At 48 h after lipofection, cells were labelled with [35 S]methionine for 30 min and chased in culture medium for the times indicated. Medium samples were removed and saved and the cells were harvested in immunoprecipitation buffer. Samples were immunoprecipitated with anti-PC2 antibody, analysed on an SDS/9%-PAGE gel and detected by fluorography. The track marked M shows molecular mass marker proteins (kDa). (B) Glycosidase digestion of secreted hPC2 immunoreactivity. Immunoprecipitated protein from medium after an 8 h chase was treated with P.N.Gase F (P), Endo H (E) or left untreated (U). Samples were analysed on an SDS/9%-PAGE gel and detected by fluorography. The track marked M shows molecular mass marker proteins (kDa).

introduce specific point or deletion mutations within a human PC2 cDNA, and to transfect these constructs into the monkey kidney cell line COS-7. These cells were used because of the high-level expression that could be attained using the pCR expression vector.

The hPC2-constructs used in this study are shown in Figure 1. hPC2 is a 613-amino-acid protein containing: (1) a signal peptide; (2) a propeptide (amino acids 1–84) which contains two cleavage sites at Lys-Arg-Arg-Arg⁵⁶ and Arg-Lys-Lys-Arg⁸⁴; (3) a catalytic domain containing the Asp-His-Ser catalytic triad which is characteristic of serine proteases; (4) a P-domain (amino acids 415–569) which is conserved among members of the family; and (5) a C-terminal hydrophobic sequence (amino acids 592–613). Two mutants were made which from previous studies [24] were known to prevent proPC2 processing, i.e. Δ RKKR⁸⁴-hPC2 and D142N-hPC2. Further constructs were designed to investigate the role of the propeptide in PC2 secretion. However, preliminary experiments indicated that complete deletion of the propeptide affected the ability of the nascent polypeptide to fold correctly, with the result that the expressed protein was degraded within the endoplasmic reticulum (ER) (K. I. J. Shennan and K. Docherty, unpublished work). We therefore made a series of smaller deletions, one of amino acids 5–49, a second of amino acids 54–77, a third of amino acids 66–80 and a fourth in which amino acids 73–80 were deleted. Note that the primary (Arg-Lys-Lys-Arg⁸⁴) and secondary (Lys-Arg-Arg-Arg⁵⁶) processing sites are retained within all these constructs. To investigate the role of C-terminal sequences, we generated several C-terminally truncated forms of hPC2: 414-hPC2, which lacks the P-domain; 569-hPC2, which terminates at the putative C-terminus of the P-domain; and 573-hPC2, which terminates at the end of the C-terminal region of identity between PC2 and PC1/3. We also generated KDEL-hPC2, which has an ER retention signal added to its C-terminus.

Expression of wild-type hPC2 in COS-7 cells resulted in the appearance of a single intracellular 75 kDa protein (Figure 2A) which was not present in mock-transfected cells (results not shown). This protein corresponds to proPC2. The amount of the 75 kDa proPC2 protein in the cell decreased after 2 h, coinciding with the appearance in the medium of three proteins of 75, 71 and 68 kDa. Proteins of similar sizes have been observed following expression of hPC2 in *Xenopus* oocytes [24] as well as in isolated rat islets of Langerhans [25], and have been shown to correspond to proPC2 (75 kDa), an intermediate generated

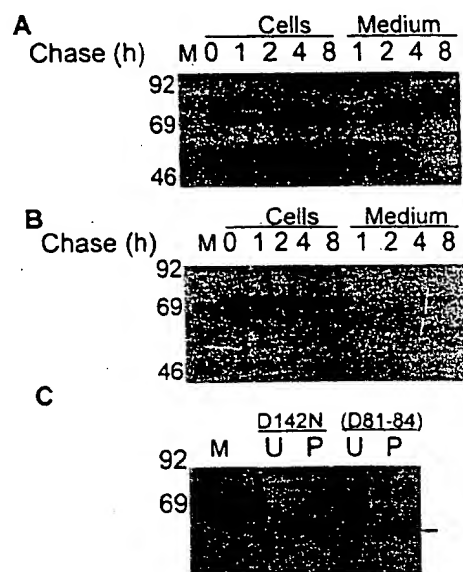


Figure 3 Expression of (Δ 81–84)hPC2 and D142N-hPC2 in COS-7 cells

Pulse-chase analysis of hPC2 immunoreactivity found in cell extracts and medium from cells transfected with pCR-(Δ 81–84)hPC2 (A) and pCR-D142N-hPC2 (B). At 48 h after lipofection, cells were labelled with [35 S]methionine for 30 min and chased in culture medium for the times indicated. Medium samples were removed and saved, and the cells were harvested in immunoprecipitation buffer. Samples were immunoprecipitated with anti-PC2 antibody, analysed on an SDS/9%-PAGE gel and detected by fluorography. The track marked M shows molecular mass marker proteins (kDa). (C) Glycosidase digestion of retained hPC2 immunoreactivity. Immunoprecipitated protein from extracts of cells transfected with either (Δ 81–84)hPC2 or D142N-hPC2 after an 8 h chase was treated with P.N.Gase F (P) or left untreated (U). Samples were analysed on an SDS/9%-PAGE gel and detected by fluorography. The track marked M shows molecular mass marker proteins.

following cleavage at the sequence Lys-Arg-Arg-Arg⁵⁶ (71 kDa) and mature PC2 generated following cleavage at the sequence Arg-Lys-Lys-Arg⁸⁴ (68 kDa). The 75, 71 and 68 kDa proteins appeared simultaneously in the medium, and there was no further processing in the medium during a subsequent 24 h chase period. Although the majority of the secreted material was unprocessed, these results suggest that a significant amount of processing was occurring within the COS-7 cells, and that the processed material was rapidly secreted with no detectable intracellular accumulation. Expression of hPC2 at lower levels resulted in the secretion of a higher proportion of processed forms, indicating that the processing activity is saturable and suggesting that it is due to an endogenous activity in COS-7 cells. It would not be expected that material should accumulate in the secretory pathway in this cell line, as COS-7 cells do not have a regulated secretory pathway.

Treatment with Endo H resulted in a small reduction of about 2–3 kDa in the molecular masses of all three of the secreted proteins (Figure 2B). Treatment with P.N.Gase F generated a larger decrease in molecular mass, of about 6–7 kDa. These results indicate that, although most of the asparagine-linked sugar moieties have been modified to acquire Endo H resistance, a significant fraction have not, or have undergone further modifications to re-acquire sensitivity. A 46 kDa immunoreactive protein secreted with PC2 is visible in Figure 2(B). This protein remains uncharacterized but, given the similar effects of Endo H and P.N.Gase F on this protein, it is possible that it is a C-terminal cleavage product of PC2, presumably containing the three potential glycosylation sites.

THIS PAGE BLANK (USPTO)

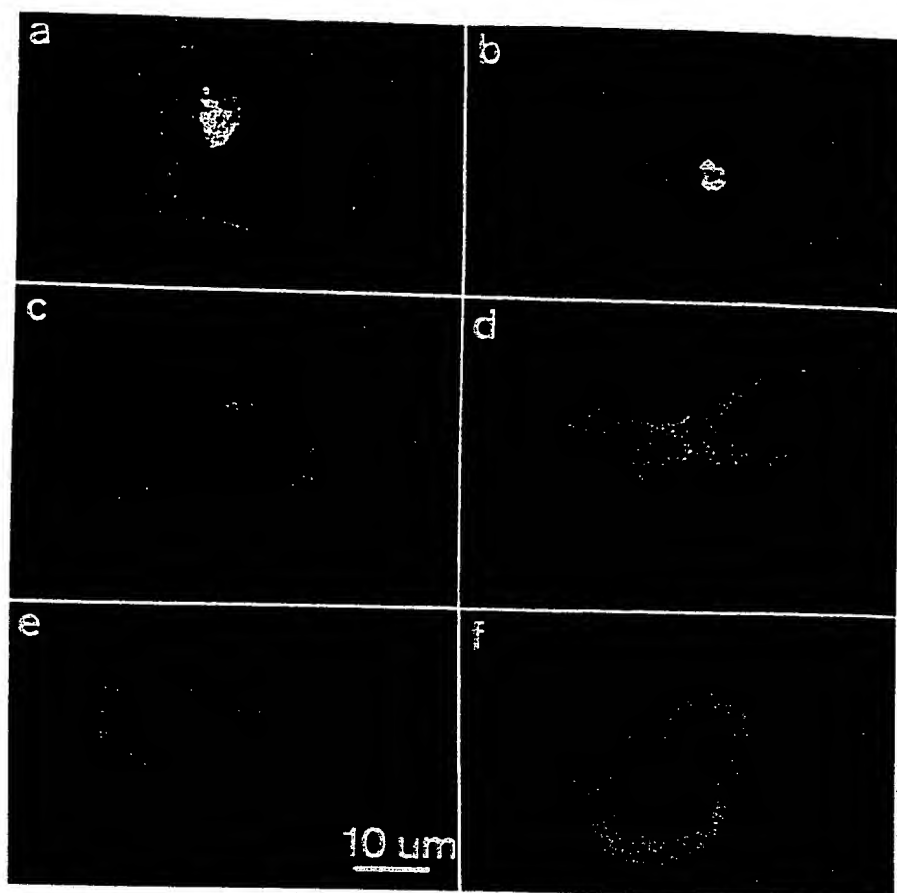


Figure 6 Immunocytochemistry of hPC2, D142N-hPC2 and KDEL-hPC2 in COS-7 cells

Cells were transfected with pCR-PC2 (a, b), pCR-D142N-PC2 (c, d) or pCR-PC2-KDEL (e, f), and plated out on to glass coverslips 48 h after lipofection and stained using the PEP4 anti-PC2 antibody and a fluorescently labelled goat anti-rabbit antibody. Staining for hPC2 is particularly concentrated in a perinuclear region typical of Golgi proteins, whereas staining for D142N-hPC2 and KDEL-hPC2 is reticular, suggesting an ER localization.

had been deleted, and so the protein could not be glycosylated. However, unlike the other mutants, 414-hPC2 appeared to be degraded in the cell, albeit slowly. Two smaller C-terminal deletions, 569-hPC2 (Figure 5B) and 573-hPC2 (Figure 5C), both of which leave the P-domain intact, prevented secretion. In both of these cases the retained protein was not degraded and underwent slow additional glycosylation.

With mutant KDEL-hPC2 (Figure 5D), a single protein of 75 kDa accumulated in the cell, with no increase in the molecular mass of the retained material. However, KDEL-hPC2 appears to be less stable in the cell, with most of the labelled protein being degraded after a 4 h chase and no labelled intracellular immunoreactivity evident after an 8 h chase.

The pulse-chase experiments identified two types of mutation, both of which resulted in the intracellular retention of the expressed protein. In one type the expressed protein was not proteolytically cleaved and the resultant polypeptide was stably retained and increased in size by about 3 kDa due to additional glycosylation. In the second type the expressed protein (i.e. 414-hPC2 and KDEL-hPC2) was also not cleaved, but did not increase in mass and was degraded over the chase period. To determine whether the two types of mutation affected the intracellular location of the expressed proteins, immunocytochemistry was performed on selected constructs. Cells expressing wild-type hPC2 exhibited the reticular pattern characteristic of

the ER plus perinuclear concentration typical of Golgi proteins (Figures 6a and 6b). Cells expressing either type of mutation, i.e. D142N-hPC2 (Figures 6c and 6d) and KDEL-hPC2 (Figures 6e and 6f) exhibited reticular staining, but with no perinuclear concentration. No significant difference in localization between the two classes of mutation could be seen at the resolution afforded by light microscopy.

DISCUSSION

Expression of hPC2 in COS-7 cells resulted predominantly in the secretion of 75 kDa proPC2, with lesser amounts of the 71 and 68 kDa processed forms. We have previously shown that *Xenopus* oocytes secrete the 75 and 71 kDa molecules and that subsequent cleavage to the 71 and 68 kDa forms occurs within the medium [24]. The 75 to 71 kDa cleavage is probably mediated by a *Xenopus* furin-like protease(s), whereas the 75/71 to 68 kDa cleavage occurs by way of an autocatalytic mechanism which is activated in the medium as it becomes slightly acidified during extended culture of the oocytes [24]. In COS-7 cells the degree of processing is not proportional to the level of expression, and in the culture medium there is no further cleavage of the 75/71 kDa species to the 68 kDa mature enzyme, suggesting that cleavage of the 75 kDa proPC2 molecule to the 71 kDa intermediate and the

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

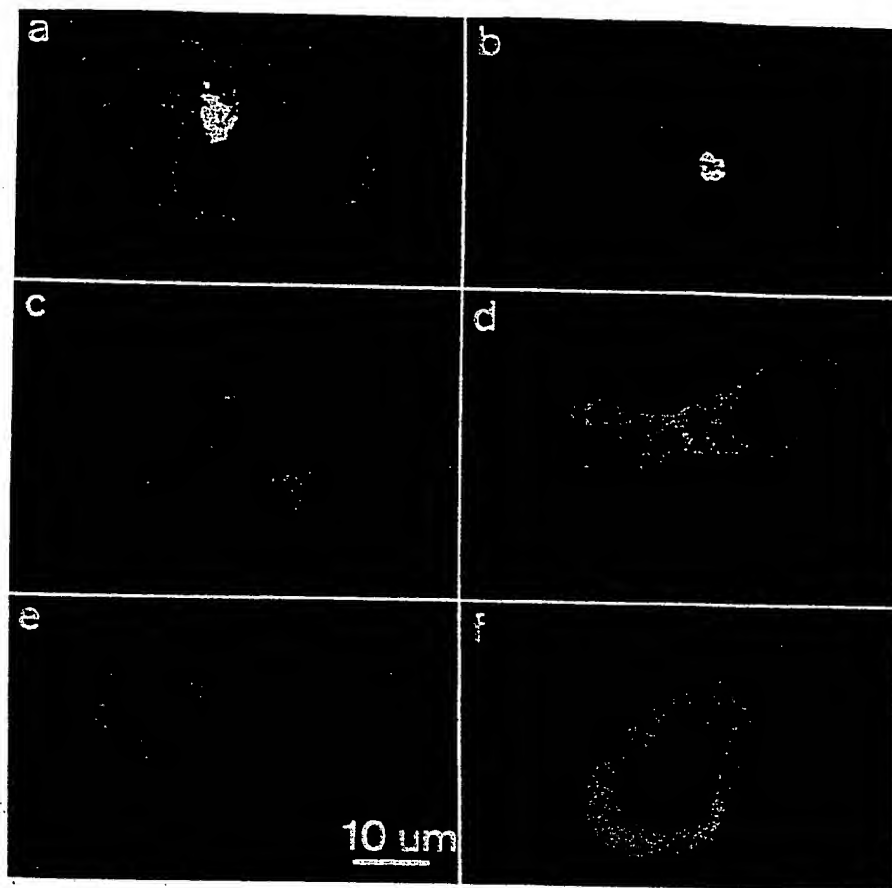


Figure 6 Immunocytochemistry of hPC2, D142N-hPC2 and KDEL-hPC2 in COS-7 cells

Cells were transfected with pCR-PC2 (a, b), pCR-D142N-PC2 (c, d) or pCR-PC2-KDEL (e, f), and plated out on to glass coverslips 48 h after lipofection and stained using the PEP4 anti-PC2 antibody and a fluorescently labelled goat anti-rabbit antibody. Staining for hPC2 is particularly concentrated in a perinuclear region typical of Golgi proteins, whereas staining for D142N-hPC2 and KDEL-hPC2 is reticular, suggesting an ER localization.

had been deleted, and so the protein could not be glycosylated. However, unlike the other mutants, 414-hPC2 appeared to be degraded in the cell, albeit slowly. Two smaller C-terminal deletions, 569-hPC2 (Figure 5B) and 573-hPC2 (Figure 5C), both of which leave the P-domain intact, prevented secretion. In both of these cases the retained protein was not degraded and underwent slow additional glycosylation.

With mutant KDEL-hPC2 (Figure 5D), a single protein of 75 kDa accumulated in the cell, with no increase in the molecular mass of the retained material. However, KDEL-hPC2 appears to be less stable in the cell, with most of the labelled protein being degraded after a 4 h chase and no labelled intracellular immunoreactivity evident after an 8 h chase.

The pulse-chase experiments identified two types of mutation, both of which resulted in the intracellular retention of the expressed protein. In one type the expressed protein was not proteolytically cleaved and the resultant polypeptide was stably retained and increased in size by about 3 kDa due to additional glycosylation. In the second type the expressed protein (i.e. 414-hPC2 and KDEL-hPC2) was also not cleaved, but did not increase in mass and was degraded over the chase period. To determine whether the two types of mutation affected the intracellular location of the expressed proteins, immunocytochemistry was performed on selected constructs. Cells expressing wild-type hPC2 exhibited the reticular pattern characteristic of

the ER plus perinuclear concentration typical of Golgi proteins (Figures 6a and 6b). Cells expressing either type of mutation, i.e. D142N-hPC2 (Figures 6c and 6d) and KDEL-hPC2 (Figures 6e and 6f) exhibited reticular staining, but with no perinuclear concentration. No significant difference in localization between the two classes of mutation could be seen at the resolution afforded by light microscopy.

DISCUSSION

Expression of hPC2 in COS-7 cells resulted predominantly in the secretion of 75 kDa proPC2, with lesser amounts of the 71 and 68 kDa processed forms. We have previously shown that *Xenopus* oocytes secrete the 75 and 71 kDa molecules and that subsequent cleavage to the 71 and 68 kDa forms occurs within the medium [24]. The 75 to 71 kDa cleavage is probably mediated by a *Xenopus* furin-like protease(s), whereas the 75/71 to 68 kDa cleavage occurs by way of an autocatalytic mechanism which is activated in the medium as it becomes slightly acidified during extended culture of the oocytes [24]. In COS-7 cells the degree of processing is not proportional to the level of expression, and in the culture medium there is no further cleavage of the 75/71 kDa species to the 68 kDa mature enzyme, suggesting that cleavage of the 75 kDa proPC2 molecule to the 71 kDa intermediate and the

THIS PAGE BLANK (USPTO)

68 kDa mature enzyme occurs within the cell and is catalysed, at low efficiency, by an endogenous protease.

The study with PNGase F indicated that all three potential glycosylation sites are utilized. Approx. 30% of the glycosyl residues on the secreted proteins are sensitive to Endo H, suggesting either that they have undergone further unusual modifications or that, for some reason, the high-molecular-mass mannose chains have not been removed. This pattern of partial Endo H sensitivity is similar to that found with PC2 immunoprecipitated from rat islets of Langerhans [25], but different from that found in transfected CHO cells or RIN msf cells, where none of the sugar residues are resistant to Endo H [26].

Mutagenesis was undertaken to investigate the structural requirements for the transfer of PC2 through the secretory pathway. Previous studies on furin had shown that cleavage of the propeptide may be a prerequisite for exit from the ER [27–29]. Dissociation of the propeptide and activation of the enzyme may occur in a later compartment. Unexpectedly, in the present study all the mutant PC2 molecules were retained within the cell as a proenzyme. Thus, whereas wild-type proPC2 (75 kDa) was efficiently secreted from COS-7 cells, mutations within the catalytic pocket or cleavage site, or deletion of propeptide sequences flanking the cleavage site or of sequences at the C-terminus of proPC2, resulted in the retention of proPC2 within the cell. These results suggest that the efficient transfer of proPC2 through the secretory pathway is dependent on a structural conformation or post-translational modification that is disrupted by disparate mutations spanning the length of the protein.

Homology modelling of furin suggests that residues +1 to –6 are important in interactions with the substrate-binding region [30]. As PC2 shows substantial sequence identity with furin, it seems likely that these interactions are also important for substrate recognition in PC2, and it is possible that some structural motif between residues 73 and 77 (i.e. –8 to –12) is required. Recent molecular modelling studies have suggested that the primary cleavage site folds into the active site of the prohormone, and it may be that the deletions in the propeptide [i.e. (Δ 54–77)hPC2, (Δ 66–80)hPC2 and (Δ 73–80)hPC2], or deletion of the Arg-Lys-Lys-Arg⁸⁴ cleavage site, or mutagenesis of the catalytic pocket (D142N), may affect this interaction, and that correct insertion of the propeptide into the active site is required before exit from the ER can occur.

There may be a different explanation for the retention of the (Δ 5–49)hPC2 mutant. This deletion causes an approx. 3-fold reduction in the rate of initial glycosylation, and this mutant does not acquire the additional glycosylation found with the other propeptide deletion mutants. It is likely that the decrease in the glycosylation rate is a result of misfolding of the molecule, since it is assumed that this region of the protein acts as an intramolecular chaperone, as is the case with the homologous bacterial protease subtilisin BPN' [31]. In subtilisin a portion of the propeptide is required for formation of the correct tertiary structure, and denatured protein will only refold if the propeptide is present, whether in *cis* or *trans* conformation.

The C-terminally truncated mutant 414-hPC2 is turned over in the cell relatively rapidly. In the case of Kex2 it was found that mutant proteins containing deletions in the P-domain were catalytically inactive and were not secreted. It was suggested that they may be acting by disrupting recognition of the processing site, and that the mutants accumulate in the cell because of their inability to undergo processing to remove the N-terminal propeptide [17]. However, in the case of PC2 it appears that deletion of the P-domain causes a structural alteration that results in degradation of the resultant proenzyme in the ER. In contrast,

the other two C-terminally truncated mutants, 569-hPC2 and 573-hPC2, behaved in an identical manner to D142N-hPC2 and the four peptides containing deletions adjacent to and including the primary processing site. This is puzzling, as neither of these deletions stretch into the P-domain, as defined by identity with Kex2, and we have seen that a smaller deletion in the C-terminus (PC2 M1) is secreted in a similar manner to the wild type [24]. The structures of the P-domain and the C-terminus are unknown at present, and it may be that these regions interact with the propeptide in some way.

Pulse-chase analysis of cells expressing KDEL-hPC2 showed that, as expected, this mutant is not secreted. The protein does not undergo the additional glycosylation found with most of the mutants and is degraded relatively rapidly. The similarity of the fate of this mutant to that of 414-hPC2 suggests that both of these proteins are retained in the same compartment.

The similarity in the immunocytochemistry for KDEL-hPC2 and D142N-hPC2 suggests that all the mutants are retained in the ER. However, the differences in additional glycosylation suggest that the majority of the mutants are retained less efficiently (and possibly therefore by a different mechanism) than KDEL-hPC2. It does not appear that these mutants are simply trapped in the ER, or are retained by the same retrieval pathway as KDEL-hPC2. It is possible that these mutants are blocked in a very early Golgi compartment and 'back up' into the ER.

This work was supported by a grant from the MRC. We thank Dr. Don Steiner for providing the PEP4 rabbit anti-PC2 antibody and the hPC2 cDNA.

REFERENCES

1. Smeekens, S. P. and Steiner, D. F. S. (1990) *J. Biol. Chem.* **265**, 2997–3006.
2. Seidah, N. G., Gaspar, L., Mien, P., Marcinkiewicz, M., Mbikay, M. and Chrétien, M. (1990) *DNA Cell Biol.* **9**, 415–424.
3. Fuller, R. S., Sterne, R. E. and Thorner, J. (1988) *Annu. Rev. Physiol.* **50**, 345–362.
4. Mizuno, K., Nakamura, T., Ohshima, T., Tanaka, S. and Matsuo, H. (1988) *Biochem. Biophys. Res. Commun.* **156**, 246–254.
5. Davey, J., Davis, K., Imai, Y., Yamamoto, M. and Matthews, G. (1994) *EMBO J.* **13**, 5901–5921.
6. Van den Ouweland, A. M. W., Duijnhoven, H. P., Keizer, G. D., Dossers, L. C. J. and Van de Ven, W. J. M. (1990) *Nucleic Acids Res.* **18**, 664.
7. Keiler, M. C., Tucker, J. E., Joh, R., Landsberg, K. E., Saltman, D. and Barr, P. J. (1991) *DNA Cell Biol.* **10**, 757–769.
8. Smeekens, S. P., Avruch, A. S., LaMendola, J., Chan, S. J. and Steiner, D. F. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 340–344.
9. Seidah, N. G., Marcinkiewicz, M., Benjannet, S., Gaspar, L., Beaubien, G., Mattei, M. G., Lazure, C., Mbikay, M. and Chrétien, M. (1991) *Mol. Endocrinol.* **5**, 111–122.
10. Nakayama, K., Hosaka, M., Hatazawa, K. and Murakami, K. (1991) *J. Biochem. (Tokyo)* **109**, 803–806.
11. Nakagawa, T., Hosaka, M., Torii, S., Watanabe, T., Murakami, K. and Nakayama, K. (1993) *J. Biochem. (Tokyo)* **113**, 132–135.
12. Meerabux, J., Yaspo, M. L., Roebroek, A. J., Van de Ven, W. J., Lister, T. A. and Young, B. D. (1996) *Cancer Res.* **56**, 448–451.
13. Seidah, N. G., Hamelin, J., Mamabachi, M., Dong, W., Tadros, H., Mbikay, M., Chrétien, M. and Day, R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3388–3393.
14. Bruzzaniti, A., Goodge, K., Jay, P., Taviaux, S. A., Lam, M. H. C., Berta, P., Martin, T. J., Moseley, J. M. and Gillespie, M. T. (1996) *Biochem. J.* **314**, 727–731.
15. Fuller, R. S., Brenner, C., Gluschkof, P. and Wilcox, R. S. (1991) in *Advances in Life Sciences: Methods in Protein Sequence Analysis* (Jörnvall, H., Höög, J.-O. and Gustavsson, A.-M., eds.), pp. 205–214. Birkhäuser, Berlin.
16. Creemers, J. M. W., Seizen, R. J., Roebroek, A. J. M., Ayoubi, T. A. Y., Huylebroeck, D. and Van de Ven, W. J. M. (1993) *J. Biol. Chem.* **268**, 21826–21834.
17. Gluschkof, P. and Fuller, R. (1994) *EMBO J.* **13**, 2280–2288.
18. Matthews, G., Shennan, K. I. J. S., Seal, A. J., Taylor, N. A., Colman, A. and Docherty, K. (1994) *J. Biol. Chem.* **269**, 588–592.
19. Shennan, K. I. J. S., Taylor, N. A., Germany, J., Matthews, G. and Docherty, K. (1995) *J. Biol. Chem.* **270**, 1402–1407.
20. Zhu, X., Ohta, Y., Jordan, F. and Inouye, M. (1989) *Nature (London)* **339**, 483–484.

THIS PAGE BLANK (USPTO)

- 21 Shennan, K. I. J. S., Taylor, N. A. and Docherty, K. (1994) *J. Biol. Chem.* **269**, 18646–18650
- 22 Campbell, M. J. (1995) *BioTechniques* **18**, 1027–1032
- 23 Cramer, L. P. and Cutler, D. F. (1992) in *Protein Targeting: A Practical Approach* (Magee, A. I. and Wileman, T., eds.), pp. 59–86, IRL Press, Oxford
- 24 Shennan, K. I. J. S., Seal, A. J., Smeekens, S., Steiner, D. F. S. and Docherty, K. (1991) *J. Biol. Chem.* **266**, 24011–24017
- 25 Guest, P. C., Arden, S. D., Bennett, D. L., Clarke, A., Rutherford, N. G. and Hutton, J. C. (1992) *J. Biol. Chem.* **267**, 22401–22406
- 26 Shen, F.-S., Seidah, N. G. and Lindberg, I. (1993) *J. Biol. Chem.* **268**, 24910–24915
- 27 Creemers, J. M. W., Vey, M., Schäfer, W., Ayoubi, T. A. Y., Roebroek, A. J. M., Klenk, H.-D., Garten, W. and Van de Ven, W. J. M. (1995) *J. Biol. Chem.* **270**, 2695–2702
- 28 Molloy, S. S., Thomas, L., VanSlyke, J. K., Sternberg, P. E. and Thomas, G. (1994) *EMBO J.* **13**, 18–33
- 29 Vey, M., Schäfer, W., Berghöfer, S., Klenk, H.-D. and Garten, W. (1994) *J. Cell Biol.* **127**, 1829–1842
- 30 Siezen, R. J., Creemers, J. W. M. and Van de Ven, W. J. M. (1994) *Eur. J. Biochem.* **222**, 255–266
- 31 Eder, J., Rheinhecker, M. and Fehrst, A. (1993) *J. Mol. Biol.* **223**, 293–304

Received 23 July 1996; accepted 9 September 1996

THIS PAGE BLANK (USPTO)

Human Coagulation Factor X Deficiency Caused by a Mutant Signal Peptide That Blocks Cleavage by Signal Peptidase but Not Targeting and Translocation to the Endoplasmic Reticulum*

(Received for publication, July 7, 1992)

Marco Racchi†, Herbert H. Watzke‡, Katherine A. High§, and Mark O. Lively†¶

From the †Department of Biochemistry, Bowman Gray School of Medicine, Wake Forest University, Winston-Salem, North Carolina 27157 and the §Departments of Medicine and Pathology, The University of North Carolina, Chapel Hill, North Carolina 27599

Human factor X_{Santo Domingo} is a form of coagulation factor X in which a mutation within the signal peptide region of the precursor protein has been correlated genetically with a severe deficiency of factor X in the affected individual. A point mutation results in substitution of Arg for Gly at the critical -3 position of the factor X signal peptide. To determine the biochemical effect of this mutation on the biosynthesis of factor X, the wild-type and mutant factor X cDNAs were subcloned into a vector for transcription and translation *in vitro*. Translation products of mRNAs encoding portions of both mutant and wild-type proteins were used in a systematic biochemical approach to evaluate directly the effect of the mutation on targeting, transport, and proteolytic processing *in vitro*. The results show that targeting and transport of factor X_{Santo Domingo} to the endoplasmic reticulum are functionally dissociated from the removal of the signal peptide by signal peptidase. Factor X_{Santo Domingo} is translocated into the endoplasmic reticulum but not processed by signal peptidase. Transient expression of the wild-type and mutant factor X in human embryonic kidney 293 cells revealed apparently normal secretion of the glycosylated two-chain form of factor X but no secretion of factor X_{Santo Domingo}. Thus, the inability of signal peptidase to cleave factor X_{Santo Domingo} is directly responsible for the absence of circulating factor X and leads to the bleeding diathesis in the affected individual.

Secreted proteins are generally synthesized as precursors having NH₂-terminal signal sequences which target nascent secretory proteins to the endoplasmic reticulum (ER)¹ and are then removed by signal peptidase (1). Although signal

peptides that target proteins to the ER have widely variable amino acid sequences, they do have three common structural features: a net positive charge at the NH₂ terminus; a central hydrophobic region; and a carboxyl-terminal region with small, nonpolar amino acids at positions -1 and -3 (signal peptides are numbered negatively from the site of cleavage toward the NH₂ terminus of the precursor) preceding the cleavage site (2). A number of mutations within, or near, signal peptides have now been described that alter the processing of human secretory proteins (3-6). Human interferon ω 1 has a signal peptide structure that presents two alternative cleavage sites for signal peptidase resulting in circulation of two forms of the protein that differ by two amino acids (7). Additionally, an insertion/deletion polymorphism has been identified in the signal peptide of the human apolipoprotein B gene that predicts two apolipoprotein B signal peptides: one that encodes a peptide of 27 residues and one that encodes a peptide of only 24 residues (8). The effect of this mutation, if any, on secretion of apolipoprotein B has not been demonstrated.

Antithrombin Dublin is an electrophoretically fast variant of antithrombin that has been shown to be the result of a mutation at the -3 position of the signal peptide in which Val is replaced by Glu (4). Individuals expressing the mutation produce a form of antithrombin in which 2 amino acids normally found at the NH₂ terminus have been removed during synthesis. It has been proposed that the Val → Glu substitution redirects the site of cleavage by signal peptidase to a bond 2 amino acid residues toward the COOH terminus of the normal protein. This mutation appears to have no direct correlation with any pathological condition (4).

Similarly, albumin Redhill is an electrophoretically slow form of human serum albumin that contains two different mutations, one of which appears to cause signal peptidase to cleave at an alternate site (3). A substitution of Cys for Arg at the penultimate position of the *pro* peptide (not the signal, or *pre*, peptide) of preproalbumin apparently creates a preferred site for cleavage by signal peptidase. It was hypothesized, but not proven, that this mutation causes signal peptidase to preferentially cleave following the newly introduced Cys, 5 residues into the propeptide. Consequently, albumin Redhill circulates with an additional Arg residue at the NH₂ terminus that would normally have been removed during processing of proalbumin. As with antithrombin Dublin, there is no disease state associated with the presence of albumin Redhill.

Mutations in human signal peptides have been correlated with defective secretion and a consequent pathological state in only two reported cases: preproparathyroid hormone (5),

* This work was supported by National Institutes of Health Grants GM32861 (to M. O. L.) and HL48322 and HL06350 (to K. A. H.). Amino acid sequence analyses were performed in the Protein Analysis Core Laboratory and oligonucleotides were synthesized in the DNA Core Laboratory, each of the Comprehensive Cancer Center of Wake Forest University supported in part by National Institutes of Health Grants CA12197 and RR-04869, as well as a grant from the North Carolina Biotechnology Center. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: Dept. of Biochemistry, Bowman Gray School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157. Tel.: 919-716-4674; Fax: 919-716-7671.

¹ The abbreviations used are: ER, endoplasmic reticulum; HOSP, hen oviduct signal peptidase; FX, coagulation factor X; FXsd, FX Santo Domingo; FXwt, FX wild-type; Chaps, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

the cells were pulse-labeled for 30 min with 0.5 mCi/ml [35 S]methionine in Dulbecco's modified Eagle's medium/F-12 medium in the presence of 10% (v/v) dialyzed fetal bovine serum and then chased for 30 min, 2 h, and 6 h. At each time point, the medium was collected and the cells were lysed in 50 mM Tris/HCl, pH 7.5, 150 mM NaCl, 1% (v/v) Triton X-100, 0.5% (w/v) sodium deoxycholate, and 0.1% (w/v) SDS. Cell lysate fractions and culture media were immunoprecipitated using rabbit anti-human FX polyclonal antibodies (Dako Corp., Santa Barbara, CA). To determine whether the translated translocated proteins were glycosylated, 5- μ l aliquots of immunoprecipitated FX proteins were treated with peptide *N*-glycosidase F (*N*-Glycanase, Genzyme Corp., Cambridge, MA) as described (15). Proteins were separated by SDS-polyacrylamide gel electrophoresis and were detected by autoradiography.

RESULTS

cDNAs encoding both prepro-FX wild-type (prepro-FXwt) and prepro-FX_{Santo Domingo} (prepro-FXsd) were subcloned into the vector pGEM-3Zf(+) to permit transcription of mRNA for translation *in vitro* (Fig. 2). Proteins produced by cell-free translation of mRNAs can be analyzed in translocation-dependent assays that reconstitute the early steps of the secretory pathway that include membrane targeting, translocation, and proteolytic processing of nascent proteins (16). The same mRNAs can be used to prepare proteins for translocation-independent assays where the fully synthesized precursor proteins are substrates for purified signal peptidase (17). Incubation of full-length secretory precursor proteins with detergent-solubilized HOSP (14) results in cleavage of the signal peptide and the processed protein product can be detected because its smaller size usually results in an increased mobility compared to the precursor when analyzed by SDS-PAGE and autoradiography. Transcription and cell-free translation of the full-length prepro-FXwt and prepro-FXsd mRNAs produced polypeptides each with an apparent molecular mass of 64 kDa based on the mobility during SDS-PAGE (under reducing conditions) although the calculated molecular mass of the polypeptide chain is 54,377 Da. These translation products were immunoprecipitable with anti-human FX antibodies. Pro-FXwt, which was expected to result from removal of the 23-residue signal peptide from full-length prepro-

FXwt after treatment with signal peptidase, was not clearly resolved from the precursor by SDS-PAGE under reducing conditions (data not shown) so an alternate approach was taken. Since it has been shown that signal peptidase normally cleaves nascent proteins during synthesis and translocation into the ER before the completion of the polypeptide chain (16), truncated versions of prepro-FX proteins were engineered so that the change in molecular weight resulting from removal of the signal peptide could be readily observed by SDS-PAGE analysis.

The first set of prepro-FX molecules with carboxyl-terminal deletions was created by transcription of prepro-FXwt and prepro-FXsd pGEM plasmids linearized by cleavage at a unique *Pst*I site within the coding region to obtain mRNAs encoding the first 127 amino acids of each protein (Fig. 2). Cell-free translation of these mRNAs yielded proteins migrating on SDS-PAGE with an apparent molecular mass of 18 kDa (Fig. 3a, lanes 1 and 3). These truncated constructs were designated prepro-FXwt/127 and prepro-FXsd/127 for

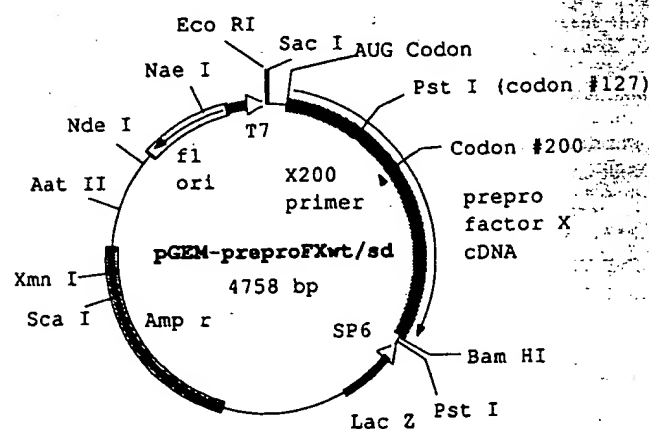


FIG. 2. Plasmid constructs of full-length and truncated forms of prepro-FXwt/sd. For the transcription of the full-length mRNAs, the plasmids were linearized with *Bam*HI endonuclease which cleaves the plasmid downstream from the termination codon. To obtain the truncated prepro-FX/127, either wild-type or mutant, the plasmids were linearized with *Pst*I endonuclease at an internal site of the FX cDNA sequence following codon 127. The construction of prepro-FX/200 (wt or sd) was accomplished by polymerase chain reaction amplification of a 712-base pair fragment from the plasmids, which includes the first 200 codons of the FX sequence, as described under "Experimental Procedures." The location of the X200 primer is shown.

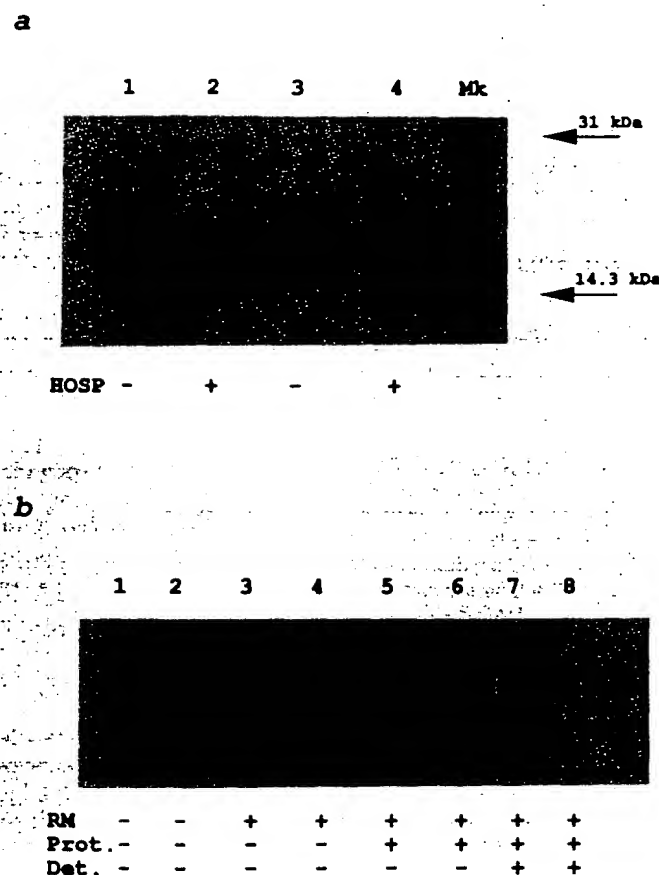


FIG. 3. Assays for processing, translocation, and segregation of prepro-FXwt/sd. a, translocation-independent processing assays (17) were performed on truncated prepro-FX/127 (wt and sd) obtained by translation of mRNAs in a wheat germ extract system supplemented with 1 mCi/ml [3 H]Leu. Aliquots of the translation mixture were subsequently incubated with HOSP (+) or buffer only (-) as described under "Experimental Procedures." Lanes 1 and 2 show the translation products from the wild-type constructs while lanes 3 and 4 show the mutant proteins. b, for targeting, translocation, and segregation assays, the mRNAs coding for prepro-FX/200 (wt and sd) were translated in a rabbit reticulocyte lysate system supplemented with 1 mCi/ml [35 S]Met in the presence (+) or absence (-) of dog pancreas rough microsomes (RM). After translation was completed, aliquots of the translation reactions were treated with a mixture of trypsin and chymotrypsin (Prot.) in the presence (+) or absence (-) of 1% (v/v) Triton X-100 (Det.). Lanes 1, 3, 5, and 7 show prepro-FXwt/200; lanes 2, 4, 6, and 8 show the mutant protein, prepro-FXsd/200.

THIS PAGE BLANK (USPTO)

the normal and mutant proteins, respectively. The products of cell-free translation were then treated with purified HOSP in a translocation-independent assay to determine the ability of the peptidase to cleave each truncated precursor. HOSP cleaved prepro-FXwt/127 to yield a faster migrating protein band which was consistent with the removal of the signal peptide (Fig. 3a, lane 2). However, HOSP was unable to cleave prepro-FXsd/127 (Fig. 3a, lane 4) which differs only by the presence of Arg in place of Gly at the -3 position before the predicted (see below) signal peptidase cleavage site.

To directly determine the site of cleavage by HOSP, prepro-FXwt/127 labeled with either [^3H]Ile or [^3H]Ala was digested with HOSP. The product of signal peptidase cleavage (Fig. 3a, lane 2) was isolated and subjected to automated amino acid sequence analysis as described under "Experimental Procedures." Sequence analysis of the signal peptidase cleavage product of [^3H]Ile-prepro-FXwt/127 released ^3H at cycles 3 and 11 (Fig. 4A), whereas the cleavage product of [^3H]Ala-prepro-FXwt/127 released ^3H at cycle 8 (Fig. 4B). These results are consistent with cleavage after Ser²³ (Fig. 1). In later cycles of sequence analysis of this protein, a larger than usual increase in the carryover of amino acid residues from one cycle to the next accounts for the peaks of radioactivity observed in cycles 12 (Fig. 4A) and 9 (Fig. 4B). Similar problems were encountered upon sequence analysis of the full-length prepro-FXwt (not shown) and are attributed to

the difficulties encountered with the sequence analysis method.

Having established that the substitution of Arg for Gly blocked translocation-independent cleavage of prepro-FXsd by signal peptidase, we next investigated the effect of the mutation on the membrane targeting and translocation functions of its signal peptide. A second set of truncated molecules, designated prepro-FX/200, was prepared containing the first 200 amino acids of prepro-FX (Fig. 2). Because there were no convenient restriction sites within the coding sequence that would yield a truncation mRNA of the desired size, the polymerase chain reaction (11) was used. Two synthetic oligonucleotides were designed to bracket a 712-base pair region including the T7 RNA polymerase promoter region and the first 200 amino acids of the prepro-FX coding sequence (see "Experimental Procedures"). These oligonucleotides were used as polymerase chain reaction primers to amplify DNA from prepro-FXwt and prepro-FXsd pGEM plasmids. The DNA amplification products were then directly transcribed by T7 RNA polymerase and the resulting mRNAs translated in the cell-free synthesis system. Translation of each truncated mRNA yielded a protein migrating with an apparent molecular mass of 24 kDa (Fig. 3b, lanes 1 and 2). As with the 127-residue FX molecules, HOSP cleaved prepro-FXwt/200 but not prepro-FXsd/200 in translocation-independent assays (not shown).

Prepro-FX/200 mRNAs were next translated in a cell-free protein synthesis system in the presence of dog pancreas rough microsomes (18) and the locations of the protein products were probed by addition of proteases. Proteins that are targeted to the ER and translocated to the interior of the microsomal vesicles in these translocation-dependent assays are protected from digestion by added proteolytic enzymes which cannot enter the vesicles (19). Prepro-FXwt/200 was cleaved by signal peptidase (Fig. 3b, lane 3) and the processed form was protected from proteolysis (Fig. 3b, lane 5). Although not cleaved by signal peptidase (Fig. 3b, lane 4), prepro-FXsd/200 was properly targeted to the microsomes because its unprocessed form was also protected from digestion by added proteases (Fig. 3b, lane 6). The protection from proteases observed in each case must have resulted from insertion of prepro-FXsd/200 and pro-FXwt/200 into the microsomes because addition of detergent to the reaction mixture allowed the proteases to penetrate the microsomes and destroy all protected FX molecules, establishing that they were not inherently stable to the protease digestion (Fig. 3b, lanes 7 and 8).

Following cell-free synthesis of prepro-FX molecules in the presence of microsomes, the vesicles were extracted with 0.1 M Na₂CO₃, pH 11.5, to determine whether the processed protein products were integrated into the lipid bilayer. Treatment of membrane vesicles at high pH solubilizes non-membrane proteins and leaves only integral membrane proteins associated with the sedimentable lipid bilayers (20). This technique revealed that the unprocessed prepro-FXsd/200 and prepro-FXwt/200 were primarily associated with the pelleted membranes (Fig. 5). In contrast to the uncleaved prepro-FXsd/200, the processed pro-FXwt/200 was released into the supernatant upon treatment with carbonate, further demonstrating that the cleaved form of the protein had been correctly targeted to the interior of the microsomes in a soluble state. We conclude that the uncleaved prepro-FXsd/200 is translocated into the microsomal vesicles where it is protected from proteolysis and remains anchored in the ER membrane via the uncleaved hydrophobic signal peptide.

Less than 10% of prepro-FXwt/200 was observed associated

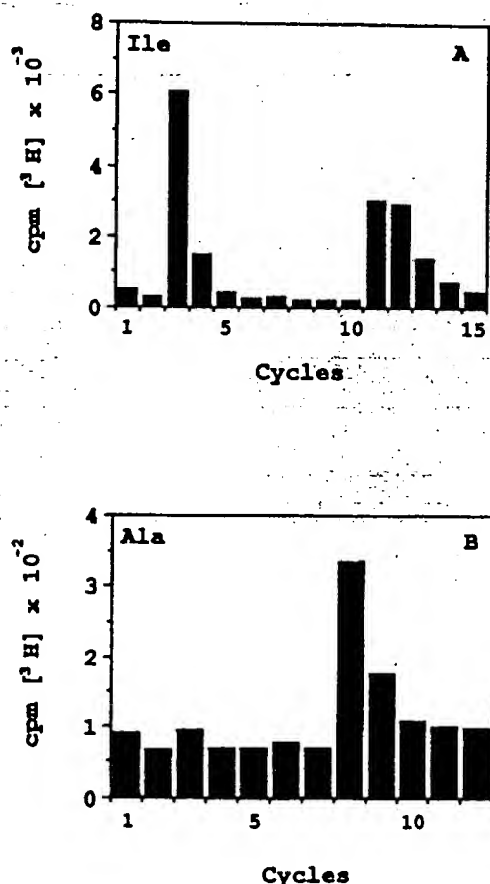


FIG. 4. Amino acid sequence analysis of the signal peptidase cleavage product of prepro-FXwt/127. The product of SP cleavage of prepro-FXwt/127 was isolated as described under "Experimental Procedures" and subjected to automated NH₂-terminal amino acid sequence analysis. Panel A shows the result of Edman degradation of pro-FX labeled with [^3H]Ile. Panel B shows the result of amino acid sequence analysis of pro-FX labeled with [^3H]Ala.

THIS PAGE BLANK (USPTO)

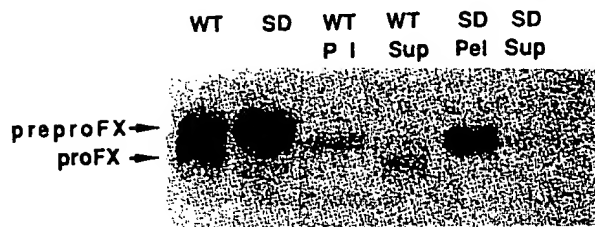


FIG. 5. Extraction of microsomes with carbonate. mRNAs coding for prepro-FX/200 wild type (WT) and Santo Domingo (SD) were translated in a rabbit reticulocyte lysate system supplemented with 1 mCi/ml [35 S]Met in the presence of dog pancreas rough microsomes. Following the completion of translation, the microsomes were treated with 0.1 M Na_2CO_3 (Ref. 10). The resulting pellet (Pel) and supernatant (Sup) fractions from each carbonate extraction were immunoprecipitated with anti-factor X then examined by SDS-polyacrylamide gel electrophoresis and autoradiography.

with the microsome pellet following carbonate extraction (Fig. 5). This precursor form must be exposed on the exterior surface of the microsomes because all uncleaved prepro-FXwt/200 molecules were shown to be susceptible to proteolysis (Fig. 3b, lane 5). This result suggests that at least a small proportion of correctly targeted precursor protein molecules are present in a membrane-bound form that is not extractable by carbonate yet remains accessible to added proteinases. This may be the result of an experimental artifact of the cell-free protein synthesis system in which the concentration of microsomes was limiting. Nevertheless, the effect of the Santo Domingo mutation on targeting of factor X is clear. In the case of prepro-FXwt/200, all translocated proteins, as defined by protection from proteolysis, are cleaved by signal peptidase. This result stands in contrast to the mutant factor X for which none of the translocated prepro-FXsd/200 molecules were cleaved.

Experiments were next designed to compare the results obtained with the truncated proteins in a cell-free system with the effect of the mutation on the secretion of the full-length proteins transiently expressed in a eukaryotic cell line. We followed the fate of pulse-labeled FX in transfected human embryonic kidney 293 cells (6), in the cell media, and in the intracellular fraction (Fig. 6). The mature, two-chain form of FXwt (21) was secreted beginning at 30 min and reached a maximum level of secretion at approximately 6 h (Fig. 6a). In contrast, FXsd was not detected in the cell medium at any time up to 24 h following the pulse of [35 S]methionine (Fig. 6b). In each case, FXwt and FXsd proteins detected in the intracellular fraction were sensitive to digestion by *N*-glycanase indicating that they were glycosylated (Fig. 7) and therefore had reached the lumen of the ER. Cleavage by *N*-glycanase results in small, measurable increases in electrophoretic mobility of the factor X proteins consistent with the removal of approximately 2 kDa of carbohydrate. These results parallel those obtained in the cell-free system, confirm the efficient translocation of prepro-FXsd into the lumen of the ER, and show the absolute block of the secretion of the mutant protein.

DISCUSSION

Signal peptides are recognized by several different proteins of the eukaryotic translocation and processing apparatus during the initial stages of targeting and transport of nascent proteins into the ER (22). They are apparently recognized first by the 54-kDa subunit of the signal recognition particle which binds to the nascent protein and directs the entire synthetic complex to the ER (23, 24). Once bound to the ER, signal peptides may also interact with components of a pos-

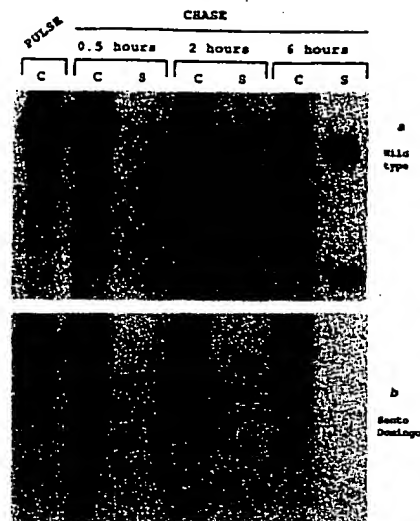


FIG. 6. Metabolic labeling of cells transfected with prepro-FXwt and with prepro-FXsd. Human embryonic kidney 293 cells transfected with full-length prepro-FX constructs were pulse-labeled for 30 min with [35 S]Met, then chased for 0.5, 2, and 6 h. Cell lysates (C) and culture media (S) were collected at each time point and immunoprecipitated using rabbit anti-human FX polyclonal antibodies. Panel a shows the results for the wild-type FX protein and panel b the results for FXsd.

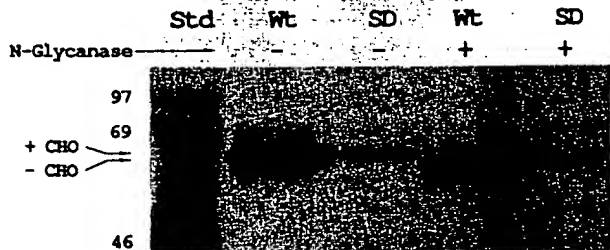


FIG. 7. Treatment of prepro-FX molecules with *N*-glycanase. Intracellular forms of factor X proteins produced during metabolic labeling in human embryonic kidney 293 cells were used to determine whether the proteins were glycosylated. The wild-type (Wt) and Santo Domingo (SD) factor X products present immediately following the chase with unlabeled methionine were either treated (+) or not treated (-) with *N*-glycanase to remove any attached carbohydrate. The products of the *N*-glycanase reaction were then separated by SDS-polyacrylamide electrophoresis and detected by autoradiography. 14 C-labeled molecular weight standards (Std) are shown in the first lane.

tulated translocation apparatus (25). A 35-kDa ER membrane protein has been identified by photochemical cross-linking as a putative signal sequence receptor (26). Additional proteins yet to be identified may interact with the peptide as it is inserted into the translocation site. Finally, most ER signal peptides are proteolytically removed by signal peptidase once the nascent protein has begun the process of transport or insertion into the ER (16). It is difficult to define those aspects of the signal peptide that serve as recognition determinants for the various proteins that bind them because the amino acid sequences of individual signal peptides vary considerably (1, 27). Because the structural characteristics that mediate the specific interactions of signal peptides with the multiple components of the translocation mechanism are not yet well understood, the occurrence of a mutation within a signal peptide can have unpredictable consequences. The experiments described here provide an approach using several established techniques to dissect the early steps of the secretory pathway and identify those affected by mutations in the signal peptide region. The use of truncated mRNAs to produce

THIS PAGE BLANK (USPTO)

shorter precursor protein molecules is especially useful in those cases where the removal of the signal peptide from larger proteins is not easily demonstrated by SDS-PAGE methods.

Since the interaction of the signal peptide with the signal recognition particle is the first critical stage in targeting of the nascent polypeptide chain to the ER, some signal peptide mutants abrogate targeting and cause the protein to be synthesized in the cytoplasm (1). Our experiments demonstrate that the targeting step is not blocked by the mutation in FXsd as the mutant protein is delivered to microsomal vesicles during cell-free protein synthesis (Fig. 3) and to the ER during synthesis in intact cells (Fig. 6). These experiments demonstrate further that the mutation also does not block the process of translocation into the lumen. It is the final step in signal peptide function, its recognition and cleavage by signal peptidase, that is dramatically impaired in the case of FXsd as a direct result of the substitution of Arg for Gly at the -3 position of the FXsd signal peptide. Both truncated forms of prepro-FXwt analyzed in these experiments were cleaved to pro-FX by detergent-solubilized signal peptidase *in vitro*. Under the same experimental conditions, cleavage of the truncated prepro-FXsd molecules was not observed.

Sequence analysis of the signal peptidase cleavage product of prepro-FXwt provided the first direct demonstration of the site of processing of human prepro-FX. A preliminary report on the site of cleavage of the signal peptide for the bovine protein has appeared (28) and the site in human prepro-FX has been correctly inferred from comparison of the gene organization of other vitamin K-dependent coagulation factors (29, 30), but this study is the first to demonstrate the site directly. Signal peptidase cleaved prepro-FX following Ser²³.

Among the general structural features that are present in all signal peptides, the amino acids close to the cleavage site appear to have the strongest influence on the specificity of cleavage by signal peptidase (2, 31). Signal peptides must have amino acids with small side chains at the -1 and -3 sites immediately before the site of cleavage (2). Large aromatic, charged, or polar amino acid residues are not observed at these sites. The placement of Arg at the critical -3 position of the prepro-FXsd mutant signal peptide clearly interferes with cleavage. It should be noted that this mutation did not induce a shift in signal peptidase cleavage site, an effect that has been observed in other cases. For example, substitution of Glu for Val at the -3 position of antithrombin Dublin results in redirection of signal peptidase cleavage to a new site 2 residues toward the COOH terminus of the protein (4). Similarly, albumin Redhill is another example in which a mutation appears to redirect signal peptidase cleavage to a new site (3). Signal peptidase appears to have some degree of flexibility in its selection of the site for cleavage if a suitable alternative site is present. In the case of prepro-FXsd, it appears that a suitable alternative cleavage site is not available so the result of the mutation is to block cleavage completely.

The results obtained in the cell-free system were further corroborated by the results of the transient expression of the full-length prepro-FX wild-type and Santo Domingo proteins in human embryonic kidney cells. The wild-type protein was processed and secreted efficiently in its two-chain zymogen form (21) while the mutant FX was never detectable in the cell culture medium. Consistent with the observations *in vitro*, the protein detected in the intracellular compartment was glycosylated demonstrating that the mutant protein was indeed translocated into the ER where glycosylation takes place.

The finding that the uncleaved signal peptide anchors prepro-FXsd/200 in the microsomal membrane suggests the possibility that uncleaved prepro-FXsd remains inserted in the ER *in vivo*. As an abnormal ER membrane protein, it is likely that prepro-FXsd is retained in the ER. All membrane or soluble proteins that are retained in the ER because of improper folding, failed oligomerization with required subunits, or aberrant post-translational processing, are eventually degraded (32, 33). Sometimes this degradation process is quite rapid. In the case of prepro-FXsd transfected in human embryonic kidney cells, the protein is relatively stable although it is evident that the protein is degraded slowly in these cells. The patient expressing this mutation did have a very low level of FX antigen detectable in her blood which could represent a small portion of prepro-FXsd that escaped degradation.

It is now clear that mutations in signal peptides of human secretory proteins can have serious consequences. Although only two examples have been described thus far which result in disease, preproparathyroid hormone (5) and factor X Santo Domingo (6), additional examples will surely be recognized in the future. As described here, a systematic approach to the study of the specific effects of such mutations on the earliest stages of the biosynthesis of secretory or membrane proteins will lead to more complete understanding of the role of signal peptides and signal peptidase in human physiology.

Acknowledgments—We thank Dr. Gregory Shelness for the gift of dog pancreas microsomes. We also thank Drs. Shelness, Fred Perrino, and Reidar Wallin for helpful discussions.

REFERENCES

- Gierasch, L. M. (1989) *Biochemistry* **28**, 923-930
- von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683-4691
- Brennan, S. O., Myles, T., Peach, R. J., Donaldson, D., and George, P. M. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 26-30
- Daly, M., Bruce, D., Perry, D. J., Price, J., Harper, P. L., O'Meara, A., and Carrell, R. W. (1990) *FEBS Lett.* **273**, 87-90
- Arnold, A., Horst, S. A., Gardella, T. J., Baba, H., Levine, M. A., and Kronenberg, H. M. (1990) *J. Clin. Invest.* **86**, 1084-1087
- Watzke, H. H., Wallmark, A., Hamaguchi, N., Giardina, P., Stafford, D. W., and High, K. A. (1991) *J. Clin. Invest.* **88**, 1685-1689
- Adolf, G. R., Maurel-Fogy, I., Kalsner, I., and Cantell, K. (1990) *J. Biol. Chem.* **265**, 9290-9295
- Visvikis, S., Chan, L., Siest, G., Drouin, P., and Boerwinkle, E. (1990) *Hum. Genet.* **84**, 373-375
- Nothwehr, S., and Gordon, J. I. (1989) *J. Biol. Chem.* **264**, 3979-3987
- Nothwehr, S. F., Hoeltzli, S. D., Allen, K. L., Lively, M. O., and Gordon, J. I. (1990) *J. Biol. Chem.* **265**, 21797-21803
- Saiki, R. K., Gelfand, D. H., Stoffel, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, M. A. (1988) *Science* **239**, 487-491
- Laemmli, U. K. (1970) *Nature* **227**, 680-685
- Towbin, H., Starbelin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 4350-4354
- Baker, R. K., and Lively, M. O. (1987) *Biochemistry* **26**, 8561-8567
- Stanton, C., and Wallin, R. (1992) *Biochem. J.* **284**, 25-31
- Blöbel, G., and Dobberstein, B. (1975) *J. Cell Biol.* **67**, 835-851
- Jackson, R. C. (1983) *Methods Enzymol.* **96**, 784-794
- Walter, P., and Blöbel, G. (1983) *Methods Enzymol.* **96**, 84-93
- Scheele, G. (1983) *Methods Enzymol.* **96**, 94-111
- Fujiki, Y., Hubbard, A. L., Fowler, S., and Lazarow, P. B. (1982) *J. Cell Biol.* **93**, 97-102
- DiScipio, R. G., Hermanson, M. A., Yates, S. G., and Davie, E. W. (1977) *Biochemistry* **16**, 698-706
- Walter, P., Gilmore, R., and Blöbel, G. (1984) *Cell* **38**, 5-8
- Kurzchalia, T. V., Wiedmann, M., Girshovich, A. S., Bochkareva, E. S., Bielka, H., and Rapoport, T. A. (1986) *Nature* **320**, 634-636
- Krieg, U. C., Walter, P., and Johnson, A. E. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 8604-8608
- Walter, P., and Lingappa, V. R. (1986) *Annu. Rev. Cell Biol.* **2**, 499-516
- Wiedmann, M., Kurzchalia, T. V., Bielka, H., and Rapoport, T. A. (1987) *J. Cell Biol.* **104**, 201-208
- von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99-105
- Blanchard, R. A., Faye, K. L. K., and Barnett, J. M. (1985) *Blood* **66** (suppl.) 331a (abstr.)
- Foster, D. C., Rudinski, M. S., Schach, B. G., Berkner, K. L., Ashok Kumar, A., Hagen, F. S., Sprecher, C. A., Insley, M. Y., and Davie, E. W. (1987) *Biochemistry* **26**, 7003-7011
- Leytus, S. P., Foster, D. C., Kurachi, K., and Davie, E. W. (1986) *Biochemistry* **25**, 5098-5102
- Perlman, D., and Halvorson, H. O. (1983) *J. Mol. Biol.* **167**, 391-409
- Klausner, R. D., and Sitia, R. (1990) *Cell* **62**, 611-614
- Bonifacio, J. S., and Lippincott-Schwartz, J. (1991) *Curr. Opin. Cell Biol.* **3**, 592-600

THIS PAGE BLANK (USPTO)

Clinical and Molecular Evidence of Abnormal Processing and Trafficking of the Vasopressin Preprohormone in a Large Kindred with Familial Neurohypophyseal Diabetes Insipidus due to A Signal Peptide Mutation*

CHARLOTTE SIGGAARD, SØREN RITTIG, THOMAS J. CORYDON,
PER HOVE ANDREASEN, THOMAS G. JENSEN, BRAGE S. ANDRESEN,
GARY L. ROBERTSON, NIELS GREGERSEN, LARS BOLUND, AND
ERLING B. PEDERSEN

Research Laboratory of Nephrology and Hypertension (C.S., E.B.P.), the Department of Pediatrics, Skejby Sygehus (S.R.), and Research Unit for Molecular Medicine (P.H.A., B.S.A., N.G.), Aarhus University Hospital, and the Institute of Human Genetics, Aarhus University (T.J.C., T.G.J., L.B.), Aarhus, Denmark; and the Department of Medicine, Northwestern University Medical School (G.L.R.), Chicago, Illinois 60611

ABSTRACT

The autosomal dominant form of familial neurohypophyseal diabetes insipidus (adFNDI) is a rare disease characterized by postnatal onset of polyuria and a deficient neurosecretion of the antidiuretic hormone, arginine vasopressin (AVP). Since 1991, adFNDI has been linked to 31 different mutations of the gene that codes for the vasopressin-neurophysin II (AVP-NP_{II}) precursor.

The aims of the present study were to relate the clinical phenotype to the specific genotype and to the molecular genetic effects of the most frequently reported adFNDI mutation located at the cleavage site of the signal peptide of AVP-NP_{II} [Ala(-1)Thr]. Genetic analysis and clinical studies of AVP secretion, urinary AVP, and urine output were performed in 16 affected and 16 unaffected family members and 11 spouses of a Danish adFNDI kindred carrying the Ala(-1)Thr mutation. Mutant complementary DNA carrying the same mutation was expressed in a neurogenic cell line (Neuro2A), and the cellular effects

were studied by Western blotting, immunocytochemistry, and AVP measurements.

The clinical studies showed a severe progressive deficiency of plasma and urinary AVP that manifested during childhood. The expression studies demonstrated that the Ala(-1)Thr mutant cells produced 8-fold less AVP than wild-type cells and accumulated excessive amounts of 23-kDa NP_{II} protein corresponding to uncleaved prepro-AVP-NP_{II}. Furthermore, a substantial portion of the intracellular AVP-NP_{II} precursor appeared to be colocalized with an endoplasmic reticulum antigen (Grp78).

These results provide independent confirmation that this Ala(-1)Thr mutation produces adFNDI by directing the production of a mutant preprohormone that accumulates in the endoplasmic reticulum, because it cannot be cleaved from the signal peptide and transported to neurosecretory vesicles for further processing and secretion. (*J Clin Endocrinol Metab* 84: 2933–2941, 1999)

THE AUTOSOMAL dominant form of familial neurohypophyseal diabetes insipidus (adFNDI) is a rare disease characterized by persistent thirst, polydipsia, polyuria, and a deficient neurosecretion of the antidiuretic hormone, arginine vasopressin (AVP) (1–5). The deficiency of AVP secretion develops early in childhood (6) and appears to be caused by degeneration of magnocellular neurons in the supraoptic and paraventricular nuclei (7–10).

Since 1991, adFNDI has been linked to 31 different mutations in 1 allele of the gene that codes for the AVP-neurophysin II precursor protein, AVP-NP_{II} (6, 11–30). This gene is located in chromosome 20 and consists of 3 exons that encode, respectively, 1) the 19-amino acid signal peptide

(SP), vasopressin (AVP), and the amino-terminal region of the transport protein NP_{II}; 2) the highly conserved central part of NP_{II}; and 3) the carboxyl-terminal region of NP_{II} and a glycoprotein, copeptin (31). Of the 31 different mutations identified in adFNDI, 4 are predicted to alter the signal peptide (6, 13, 17, 20–22, 24, 25), 1 is predicted to alter the AVP moiety (26), and the other 26 are predicted to alter the NP moiety (11, 12, 14–20, 23–25, 27–30). Of the 4 SP mutations, 3 are located near the cleavage site. One that substitutes threonine for alanine at position -1 is the most common mutation described in adFNDI, as it has now been found in 7 apparently unrelated families in America, Japan, Denmark, and Spain.

Based on the location and type of mutations identified in adFNDI and the lack of any major differences in the clinical phenotypes, we and others have postulated that all of the mutations act by replacing or deleting one or more amino acids important for proper folding and processing of the preprohormone (17). As a result of this defect, the mutant precursor cannot be folded and dimerized and is retained in the endoplasmic reticulum, where it accumulates and even-

Received August 31, 1998. Revision received April 12, 1999. Accepted April 19, 1999.

Address all correspondence and requests for reprints to: Charlotte Siggaard, M.D., Department of Pediatrics, Skejby Sygehus, Aarhus University Hospital, DK-8200 Aarhus N, Denmark. E-mail: rittig@iekf.au.dk.

* This work was supported by grants from the Danish Medical Research Council, the Novo Nordic Foundation, the Lundbeck Foundation, the Karen Elise Jensen Fund, and the University of Aarhus, Denmark.

THIS PAGE BLANK (USPTO)

tually kills the cell by interfering with the orderly processing of other essential proteins. This hypothesis is supported by recent *in vitro* expression studies that demonstrate accumulation of mutant AVP-NP_{II} precursor within the endoplasmic reticulum (ER) (32, 33). In the present study, we have investigated both the clinical phenotype and the cellular and biochemical effects of the AVP-NP_{II} gene mutation most often identified in adFNDI.

Subjects and Methods

Subjects

Studies were performed on 32 living members (16 men and 16 women; mean age, 34 yr; range, 2–69 yr) from 3 generations of a Danish kindred previously shown to have FNDI (4) in association with a missense mutation predicted to alter the C-terminal residue of the signal peptide (17) (Fig. 1). Also studied were 11 spouses of the affected kindred members (aged 47.1 ± 11 yr). The local ethical committee approved the study, and appropriate informed consent was obtained from all human subjects.

Clinical tests

Members of the family were questioned about the signs and symptoms of diabetes insipidus. All but two of those with a history of polyuria and polydipsia had one or more tests to confirm the diagnosis of diabetes insipidus. The tests included 1) measurements of the 24-h urine volume, urine osmolality, plasma osmolality, and plasma AVP when the subjects were untreated and on *ad libitum* fluid intake; 2) measurements of plasma osmolality, plasma vasopressin, and urine osmolality during a fluid deprivation test or infusion of hypertonic (3%) saline (34); and 3) measurement of spot urinary osmolality during treatment with therapeutic doses of 1-des-amino-D-arginine-8 AVP (DDAVP). In some affected and unaffected family members, vasopressin was also measured in spot urines collected randomly under conditions of *ad libitum* fluid intake.

Laboratory

Plasma AVP was measured by RIA, which was a modification of the method described previously (35). Before the assay procedure, C₁₈ Sep-Pak extraction was performed; a rabbit anti-AVP was used (ICM Immunochromatics, Tumba, Sweden). The minimum level of detectability was 0.5 pg/mL. The coefficients of variation were 13% (interassay) and 9% (intraassay). Urinary AVP was determined after lyophilization and reconstitution to isotonicity with distilled water, and the values are expressed per mg creatinine as previously described (6). Plasma and urine samples were measured for osmolality by freezing point depression (Advanced Cryometric Osmometer, 3C2, Advanced Instruments, Needham, MA). Plasma creatinine was measured by routine methods.

Amplification and sequencing of genomic DNA

Genomic DNA was extracted from the buffy coat of peripheral leukocytes as described previously (6). All exons of the VP_{NP}II gene were amplified separately by PCR using 30-bp primers flanking each exon. The primer sequences and locations as well as the PCR cycling conditions have been described previously (17). For restriction enzyme di-

gestion analysis, the PCR product was digested with the endonuclease Bst_{UI} following the manufacturer's instructions.

Construction of expression vectors

A pc-DNA 1 plasmid (Invitrogen, San Diego, CA) containing a 700-bp human wild-type complementary DNA (cDNA) fragment encoding the entire human AVP-NP_{II} precursor protein, pcDNA-HV2 (36), was provided by Prof. D. Richter, Eppendorf University (Hamburg, Germany). Sequencing of the plasmid revealed a deviation from the normal genomic sequence, namely a G to T substitution at position 2120 in the genomic sequence (position 579 in the cDNA sequence (36), predicting an amino acid change in the NP_{II} moiety (Gly⁵⁸Val). This discrepancy was corrected by replacing the distal part of the cDNA with a PCR-generated fragment from the genomic DNA of a control subject.

The signal peptide mutation Ala(-1)Thr was introduced by site-directed mutagenesis according to the PCR-based method described by Kuipers *et al.* (37). In the first PCR, a DNA fragment was amplified with a flanking primer in the central portion of exon 2 and a mutagenesis primer containing a G to A substitution at position 279 in the genomic sequence. The purified PCR product from the first PCR was then used as a megaprimer together with the T7 primer in a second PCR with the wild-type plasmid as a template. After digestion with *Hind*III and *Sma*I, the purified product from the second PCR was ligated back into the wild-type plasmid. To confirm that no PCR-derived errors were present, the constructed plasmid and the wild-type plasmid were checked by sequencing.

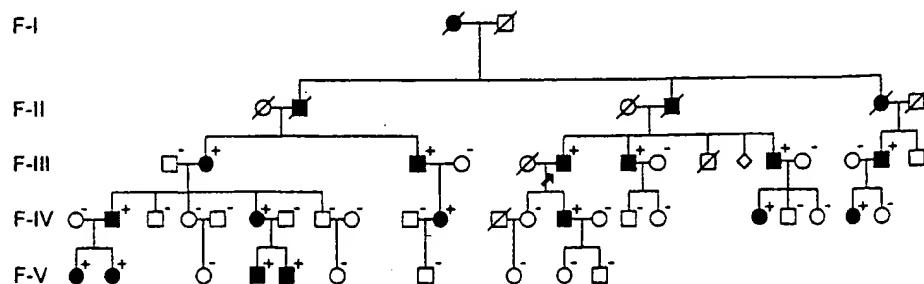
Cell culture and transfection

Mouse Neuro2A cells were obtained from the American Type Culture Collection (Manassas, VA) and were grown in DMEM (Life Technologies, Inc., Gaithersburg, MD) supplemented with 10% FCS. The cells were maintained in a 5% CO₂ atmosphere at 37°C. One day before transfection, the cells were recultivated in culture flasks at 50% confluence. Transfection of the cells with recombinant plasmids or pcDNA 1 vector as a control was performed using a calcium phosphate coprecipitation method (38, 39).

Northern blotting and measurement of secreted immunoreactive AVP

Sixteen hours after transfection, the cells were plated in either six-well plates for measurement of AVP immunoreactivity in the cell medium or in culture bottles for Northern blot analysis. The following day the cells from the six-well plates were washed in PBS and supplied with 2 mL fresh culture medium. After 24 h, the culture medium was harvested. The cells were lysed, and the protein concentration of the lysates was determined by Bradford protein assay (Bio-Rad Laboratories, Inc., Hercules, CA). For Northern blotting, the cells were grown for 48 h in culture medium. Total ribonucleic acid (RNA) was isolated from the transfected cells using a RNA extraction kit (RNAzol, WAK Chemical Co., Bad-Soden, Germany). Total RNA (5 µg) was subjected to agarose gel electrophoresis and transferred to a Zeta Probe membrane (Bio-Rad Laboratories, Inc.) by capillary blotting. A 600-bp human AVP-NP_{II} cDNA was radiolabeled (Prime It, Stratagene, La Jolla, CA) with ³²P and used as a probe. Hybridization and autoradiography were performed using standard methods.

FIG. 1. Pedigree of a Danish kindred with diagnosed adFNDI. Clinically affected subjects are marked with blackened symbols. Presence of the Ala(-1)Thr mutation by either sequencing or restriction enzyme digestion is marked with a plus. As shown, all clinically affected subjects (n = 16) and none of the unaffected (n = 27) had the mutation.



THIS PAGE BLANK (USPTO)

Western blot analysis

Forty-eight hours after transfection, cells for Western blotting were lysed with Nidet buffer (1% Nonidet P-40, 0.4% desoxycholate, 66 mmol/L ethylenediamine tetraacetate, and 10 mmol/L Tris-HCl, pH 7.4) supplemented with protease inhibitors (Trasyolol, 200 U/mL; leupeptin, 10 µg/mL; bacitracin, 200 µg/mL; phenylmethylsulfonylfluoride, 250 µmol/L; Sigma Chemical Co., St. Louis, MO) (33). When differentiated cells were studied, cells were grown in serum-free medium for 24 h before cell lysis. Cell extracts were subjected to 16.5% Tris-tricine gel electrophoresis, and separated proteins were electrotransferred to a polyvinylidene difluoride membrane using a transfer buffer containing 50 mmol/L boric acid (pH 9.0) and 20% methanol. For the detection of NPII proteins, the membranes were blocked and incubated with anti-NPII antibody (DAKO Corp., Glostrup, Denmark), and the proteins were detected with a chemiluminescence detection system (Western-Light Plus Kit, Tropix, Bedford, MA).

Immunostaining and confocal laser scanning microscopy

Immediately after transfection, cells were grown in chamber slides (Nunc, Copenhagen, Denmark) for 24 h in culture medium containing 10% serum followed by 24 h in serum-free medium to induce differentiation to neuronal cells. The cells were then fixed with 4% paraformaldehyde and immunostained essentially as described by Jensen *et al.* (40). The primary antibodies were directed specifically against NPII (anti-NPII, rabbit anti human-NPII, ICN Biochemicals, Inc. Costa Mesa, CA). The secondary antibody was fluorescein isothiocyanate-conjugated porcine rabbit antibody (Dakopatts, Copenhagen, Denmark). For confocal laser scanning microscopy (Leica Corp., Heidelberg, Germany), we used a two-layer immunostaining procedure separated by a washing period in PBS. The cells were incubated for 60 min at room temperature with the NPII antibody diluted (1:1000) in blocking reagent (Boehringer Mannheim, Mannheim, Germany) followed by incubation with the secondary antibody (1:200) for another 60 min.

To look for colocalization of NPII and endoplasmic reticulum or Golgi apparatus, cells were also incubated with either a mouse antibody directed against an ER antigen (Grp78; Stress Gen, Biotechnologies, Victoria, BC, Canada; diluted 1:50) (33) or a mouse antibody directed against a Golgi antigen (microtubule-binding Golgi membrane protein 58K, Sigma; 1:50), followed by incubation with rhodamine-conjugated goat antimouse antibody (Sigma; diluted 1:200). After washing, coverslips were mounted with one droplet of anti-fade solution (41).

The Grp78 antibody is monoclonal and identifies glucose-regulated protein 78 (Grp78) also known as Ig heavy chain binding protein and additional proteins containing the KDEL retention signal sequence. The KDEL retention signal is a carboxyl-terminal sequence (Lys-Asp-Glu-Leu) shared by luminal ER proteins and is shown to be attached to proteins retained in the ER (42). The FITC-conjugated antibody and rhodamine-conjugated antibodies are visualized by confocal laser scanning microscopy as red and green fluorescent colors, respectively. Colocalization of the NPII protein and Grp78 protein results in a merge between these two colors that is visualized as an orange/yellow fluorescent color.

Results

Clinical studies

Sixteen members (7 women and 9 men) from 3 generations of the kindred had a history of polyuria and polydipsia since childhood. The clinical results are listed in Table 1. In summary, the age of onset in this kindred was rather high and averaged 3.2 yr (range, 2.5–9 yr). All but one patient (FIII-11) showed complete diabetes insipidus with an inability to concentrate urine and a severely deficient secretion of AVP during osmotic stimulation. There was no significant gender difference in the severity of polyuria even when values were corrected for body weight (data not shown). All 15 patients responded well to treatment with DDAVP and, except for 1 who chose not to continue treatment, currently maintain normal urine volumes and osmolalities on intranasal doses ranging from 10–40 µg (mean, 22 µg/day). As in other adFNDI kindreds, 1 diseased affected family member (FI-1) had experienced a gradual decline in urine volume after the age of 50 yr to a level at which treatment was no longer necessary.

As shown in Fig. 2, mean urinary vasopressin measured in a random sample and expressed as a function of the concurrent urinary creatinine was significantly lower in 14 affected subjects than in 11 unaffected family members and 11 spouses (15 ± 8 , 39 ± 15 , and 35 ± 13 pg/mg creatinine, respectively). Although the individual values in the

TABLE 1. Clinical and hormonal data on 16 affected patients from a Danish kindred with familial neurogenic diabetes insipidus

Patient	Gender	Age ^a (yr)	BW (kg)	Age of onset (yr)	Basal				Type	Osmotic stimulation			DDAVP	
					Uvol (L/24 h)	Uosm (mosmol/kg)	Posm (mosmol/kg)	Pavp (pg/ml)		Posm _{max} (mosmol/kg)	Pavp _{max} (pg/mL)	Uosm _{max} (mosmol/kg)	DD (i.n.) (µg/day)	Uosm (mosmol/kg)
FIII-2	F	63	98		14.5	65	290	0.6	H	312	0.7	135	10	360
FIII-3	M	58	70	2–3	11.0 ^b	87							ir	
FIII-6	M	55	65	Childhood	6.8	164	293		F	310		202	20	469
FIII-7	M	51	84	Childhood	10.8	100	289		F	303		166	20	498
FIII-11	M	39	76	Childhood	6.8	222	290		F	296		343	15	473
FIII-14	M	47	82	2–3	10.0 ^b								30	561
FIV-2	M	42	128	3	28.9	71	293						40	416
FIV-6	F	38	86	3	14.7	74	296	0.5	H	311	0.7	83	40	512
FIV-11	F	33											20	902
FIV-14	M	33	88	3	11.8	168	292	0.7	H	299	0.8	196	10	546
FIV-18	F	17	55		6.6	99	286		F	292		154	20	304
FIV-21	F	22	64	9	9.0 ^b								10	733
FV-1	F	10	40	4	13.2	75	298	1.6	H	311	1.6	75	20	318
FV-2	F	8	40	2	12.0	65	294	1.1	H	302	1.3	59	40	729
FV-4	M	16	56	2	16.5	71	297	0.9	H	314	1.1	95	20	849
FV-5	M	13	41	3	12.8	63	291	0.9	H	307	1.1	85	20	641
Mean				3.4	12.4	102	292	0.9		305	1.0	144	22	554
SEM				0.6	1.4	13.9	1.0	0.14		2.2	0.13	25	2.7	48

H, Hypertonic saline infusion; F, dehydration test; DD, daily dose; ir, irregular use.

^a When tested.

^b Home measurement.

THIS PAGE BLANK (USPTO)

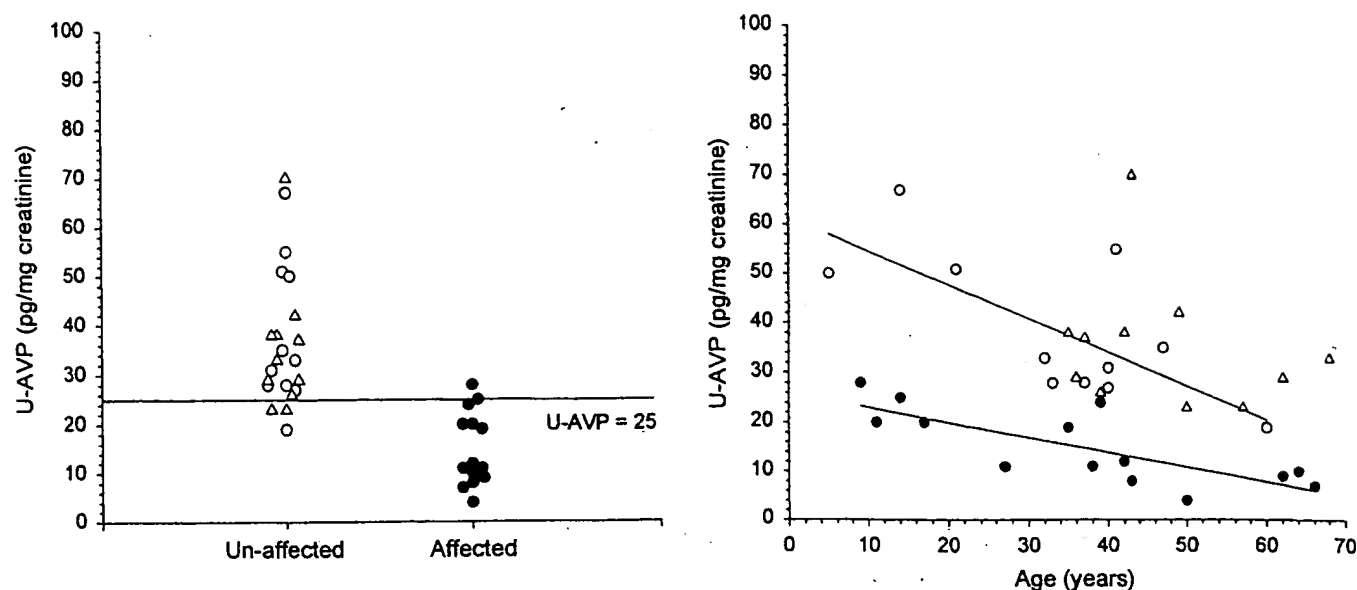


FIG. 2. The urinary excretion of vasopressin (U-AVP), as determined in spot urine and related to creatinine content, showed significantly lower values in affected individuals compared with unaffected individuals or spouses. As shown in the lower panel, an age-dependent loss of vasopressin secretory capacity was indicated by a highly significant negative correlation between U-AVP levels and age in both affected and unaffected family members. ●, Affected members; ○, unaffected members; △, spouses.

3 groups overlapped slightly, at a cut-off level of 25 pg/mg creatinine the positive predictive value for detecting adFNDI by this method was 81.3%, and the negative predictive value was 95.2%. As shown in the lower panel of Fig. 2, urinary vasopressin excretion correlated negatively with age in both affected and unaffected family members ($r = -0.78$; $P < 0.001$ and $r = -0.71$; $P < 0.05$, respectively).

Genotyping

Sequencing the coding region of AVP-NP11 gene in 12 of the 16 affected family members revealed that all of them have a single base substitution (G→A) at position 279 in exon 1 of 1 allele. This mutation predicts the exchange of Ala with Thr at the -1 position of the signal peptide moiety (Fig. 3). No abnormalities were found in the rest of the coding region of the gene. As the mutation eliminates a restriction site for the endonuclease *Bst*UI, digestion of exon 1 with this enzyme resulted in an abnormal 269-bp fragment in affected subjects (Fig. 4). This abnormal product was found in all 16 affected family members, but in none of the 16 who were unaffected or any of the 11 spouses (Figs. 1 and 4). Two children (FV-7 and FV-10) who had no symptoms but were too young to assign a phenotype both showed a normal digestion pattern.

Expression studies

As shown in Fig. 5, Neuro2A cells transiently transfected with the wild-type cDNA secreted significantly more immunoreactive AVP into the medium (41.2 ± 5.6 pg/ μ g cell protein/24 h) than the cells transfected with Ala(-1)Thr mutant cDNA (5.8 ± 0.9) or the unmodified vector (0.1 ± 0.01). Northern blotting showed similar levels of messenger RNA expression (data not shown).

As shown in Fig. 6, Western blotting of lysates from cells transiently transfected with wild-type cDNA revealed one band of protein that reacted with the NP11 antibody. It corresponded in size to the 21-kDa glycosylated pro-AVP precursor. In contrast, cells transfected with the mutant cDNA produced a slightly larger band of protein that had a molecular size of approximately 23 kDa, which is similar to uncleaved, glycosylated prepro-AVP.

As shown in Fig. 7, cells transfected with wild-type cDNA contained small aggregates of NP11 immunoreactivity (green color) that was distributed throughout the cytoplasm, especially in the tips of the cellular processes where the secretory vesicles normally are located (Fig. 7A). However, in cells expressing the mutant cDNA, the immunofluorescence was localized to the perinuclear network and was observed to a much lesser extent in the tips of the cells (Fig. 7B). Labeling with the Grp78 antibody (red color) showed a uniform perinuclear staining in both cells expressing the wild-type (Fig. 7C) and those expressing the mutant cDNA (Fig. 7D). When the cells were double labeled with the NP11 antibody (green) and the Grp78 antibody (red), the wild-type cells showed no colocalization (Fig. 7E). However, double labeling of the cells expressing the mutant cDNA showed that a substantial portion of the NP11 protein appeared to be colocalized with the ER antigen Grp78 (orange/yellow color, Fig. 7F). Double labeling of the cells with the NP11 antibody and the Golgi antibody revealed no colocalization in either wild-type or mutant cultures (not shown).

Discussion

These studies provide further clinical and molecular genetic evidence that a missense mutation affecting the -1 signal peptide residue of the AVP-NP11 preprohormone

THIS PAGE BLANK (USPTO)

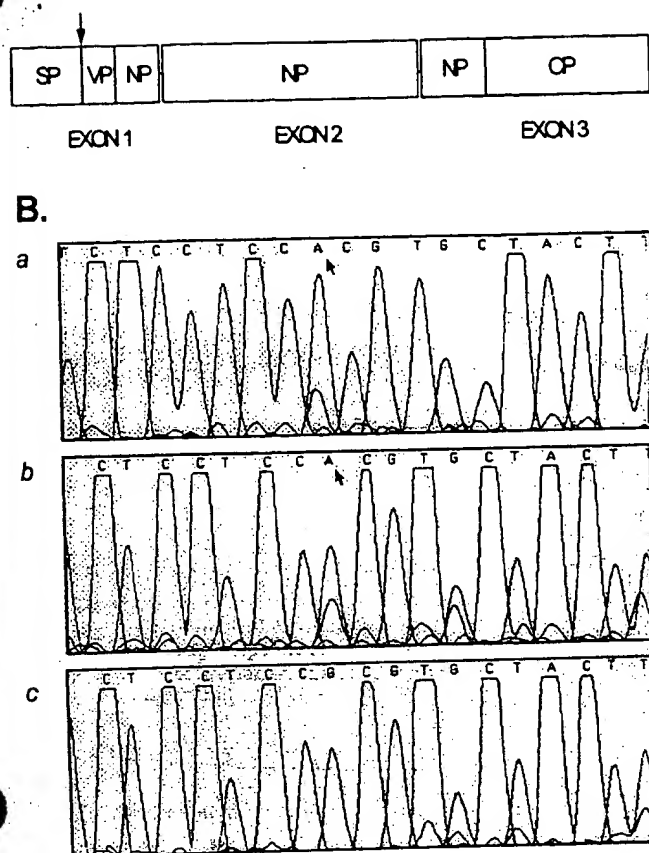


FIG. 3. A, Schematic diagram of the coding region of the AVP-NP-II gene. The arrow indicates the signal peptide mutation. B, Section of the sequencing chromatogram obtained by bidirectional dye terminator sequencing of PCR-amplified exon 1 of the AVP-NP-II gene from one of the affected members of the family (a, sense; b, antisense) and from an unaffected member (c, antisense). The arrow indicates the heterozygous mutation seen in both the sense and the antisense chromatogram from the affected family member. SP, Signal peptide; VP, vasopressin; NP, neurophysin; CP, copeptin.

causes adFNDI by directing the production of a mutant AVP-NP-II precursor that cannot be processed and routed normally. As in previous reports of kindreds with this mutation (6, 13, 17, 20, 24), it cosegregated perfectly with clinical signs of the disease: affected equal numbers of males (9) and females (7), and was transmitted to approximately 50% of those at risk. Like a previous report (6), we also found that the clinically affected members always had a greater than 50% deficiency in AVP secretion even though only one allele of the gene was mutated. These findings are fully consistent with a completely penetrant autosomal dominant trait.

The present study also supports a previous report (6) that the AVP deficiency produced by this Ala(-1)Thr mutation is not present at birth, but develops early in childhood and worsens progressively with time. Thus, we found that the DI in our kindred reportedly began at ages ranging from 2-9 yr, and the urinary AVP excretion in these patients correlated negatively with age. However, as in other adFNDI families, the debut of symptoms (2-9 yr) as well as the severity of polyuria (6.6-28.8 L/24 h) and the AVP deficiency (assessed

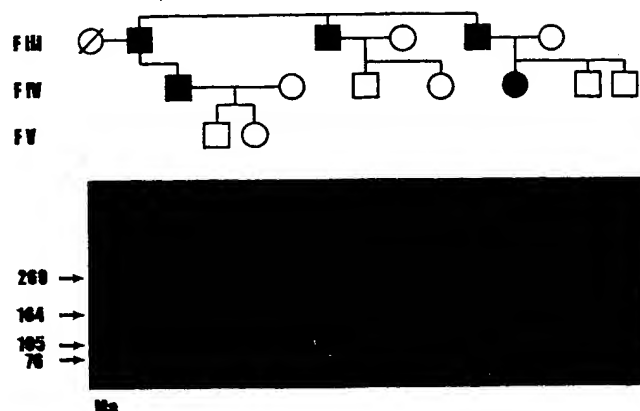


FIG. 4. *Bst*UI digestion of PCR product from genomic DNA from five affected and nine unaffected family members of the FNDI kindred. Agarose gel electrophoresis of the digestion products showed a heterozygous pattern with an abnormal band (269 bp) in all affected and in no unaffected family members.

by both plasma and urine content) varied considerably within the family. Because of this variability and the limited number of patients usually available for careful evaluation, it is difficult to determine with certainty whether there are significant differences in the severity of the disease produced by the various AVP-NP-II gene mutations. However, there is some evidence suggesting that the age of onset is lower in several kindreds with mutations in the NP-II moiety (Leu⁵⁰Pro, Gly⁵⁷Arg, and Arg⁶⁶Pro) than in those with the Ala(-1)Thr mutation in the SP (27, 29). This difference is consistent with theoretical expectations, because mutations affecting the SP cleavage site would be expected to allow the formation of some normal prohormone from the mutant alleles, whereas the NP mutations would not (17, 27).

The wide variability in the age of onset and severity of the AVP deficiency among patients with the same Ala(-1)Thr mutation is unexplained but is probably due to individual differences in other genetic or environmental influences on the neurohypophysis. For example, the rate of production of the mutant precursor could vary due to individual differences in the control of gene expression and/or the intensity of neurohypophyseal stimulation. Susceptibility to the postulated toxic effect of the mutant precursor could also vary due to individual differences in the capacity to degrade or otherwise dispose of the mutant precursor. Finally, the secretory reserve of the neurohypophysis could also vary due to individual differences in the development of the gland.

Our finding that the basal urinary AVP/creatinine ratio declines with age in unaffected as well as affected family members was unexpected because the plasma AVP response to osmotic stimulation increases with age in healthy adults (43). This discrepancy may be due to age-related differences in urinary AVP clearance or other variables, such as the rate of basal water intake and insensible loss. Either way, it is a reminder that even when adjusted to the rate of creatinine excretion, urinary AVP excretion is not always a reliable index of AVP secretion (44) or AVP secretory capacity. Thus, it raises questions about the significance of the age-related decline in urinary AVP that we and others (6) have observed in patients with adFNDI. Nevertheless, the measurement of

THIS PAGE BLANK (USPTO)

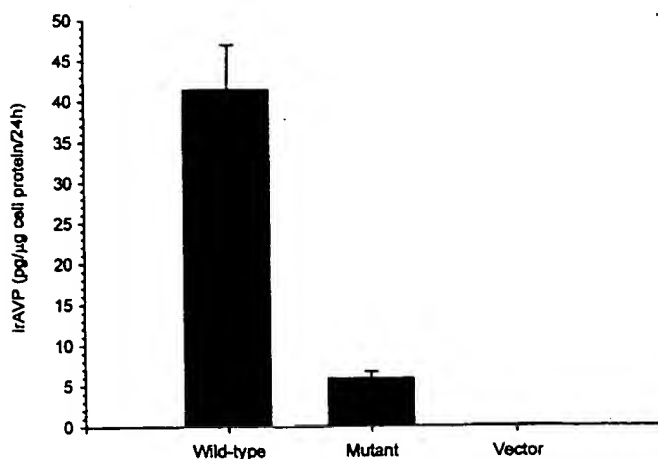


FIG. 5. RIA measurements of AVP immunoreactivity (irAVP) secreted into the cell medium for 24 h. The immunoreactivity is expressed as picograms of irAVP per μ g cell protein/24 h. The data represent the mean \pm SD from six wells and demonstrate that cells transfected with the wild-type cDNA secreted irAVP at an 8-fold higher level than cells transfected with mutant cDNA.

the AVP/creatinine ratio in spot urine samples appears to be a relatively good predictor of AVP deficiency, at least in severely affected patients. As such, it seems worthy of further investigation as an alternative to measurement of plasma AVP during osmotic stimulation.

Expression of the Ala(-1)Thr mutation in a neurogenic cell line with a secretory apparatus capable of processing the wild-type AVP-NP II gene through the regulated pathway (45, 46) provides further evidence that the mutation impairs the production of AVP by interfering with cleavage of the signal peptide from the preprohormone and its trafficking from the ER to the secretory vesicles. Thus, compared to Neuro2A cells transiently transfected with the wild-type AVP-NP II cDNA, cells transfected with the Ala(-1)Thr mutant cDNA produced about 8-fold less immunoreactive AVP, accumulated a NP II-containing protein that had electrophoretic properties indistinguishable from those of uncleaved prepro-AVP-NP II, and remained associated with the endoplasmic reticulum.

In the only strictly comparable study published to date, Ito *et al.* (32) also found that the Ala(-1)Thr mutation impaired AVP secretion by Neuro2A cells. The degree of impairment was about the same as that with two other mutations (Gly⁵⁷Ser and Cys⁶⁷stop), but appeared to be less than that with another (δ Glu⁴⁷) that was also associated with adFNDI. Robertson *et al.* also reported that Neuro2A cells transiently or stably transfected with the Ala(-1)Thr mutant produced 10- to 100-fold less immunoreactive AVP and died sooner after differentiation to postmitotic neurons than cells similarly transfected with cDNA lacking this mutation (47, 48). Although the human AVP-NP II cDNA used in these studies was found later to deviate from the standard sequence at a locus in exon 3 (see *Materials and Methods*), the relative deficiencies in AVP production and viability associated with the Ala(-1)Thr mutation are also relevant here because the deviation in exon 3 was also present in the wild-type cDNA, was located in an unconserved coding region of the gene, and

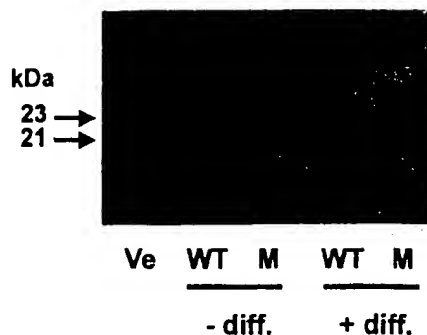


FIG. 6. Western blotting of NP II protein in the cell lysates transiently transfected with vector alone (Ve) or wild-type (WT) or mutant (M) cDNA. In both the undifferentiated and differentiated Neuro2A cells, the mutant NP II protein was larger than the wild-type signal, indicating impaired cleavage of the signal peptide. The cells were differentiated to neuronal cells by growth for 24 h in serum-free medium.

has not been reported to affect AVP production in any system.

There are several possible mechanisms by which the Ala(-1)Thr mutation could cause the observed abnormalities in vasopressin production, precursor processing, and cellular trafficking. Theoretically, the Ala(-1)Thr mutation could either impair cleavage of the signal peptide or misdirect cleavage to an alternate site one or more residues upstream (17, 49, 50). Misdirected cleavage of the signal peptide was not observed in this or another study (32) that used pulse-chase techniques, and Western blotting in the present study also revealed only one abnormal band of approximately 23 kDa, consistent with uncleaved glycosylated prepro-AVP-NP II. Whether the lack of the 21-kDa band should be taken as evidence for a complete inhibition of SP cleavage in the mutant is not clear, as it can be argued that any correctly cleaved fraction of the prohormone would be further processed into the regulated pathway and thus would not show up on a Western blot of cell lysate. In any case, failure to remove all of the signal peptide from the N-terminus of the AVP moiety is likely to hinder its insertion in the AVP-binding pocket in neurophysin, a process that is necessary for normal folding and dimerization of the prohormone (51).

The present study also confirms that, in contrast to the wild-type prepro-AVP-NP II, a substantial portion of the mutant precursor appears to be colocalized with the ER. Immunostaining using confocal laser scanning microscopy clearly demonstrates that antibodies against the mutant protein colocalize with those against the Grp78 protein that resides in the ER. In contrast, no colocalization was found using an antibody against a Golgi protein (data not shown). In contrast, antibodies against the wild-type protein did not colocalize to either the ER or Golgi, but were concentrated elsewhere in the cell body, especially in the cell processes. This finding agrees with the results of other studies (32, 33), which showed that this and other mutations associated with adFNDI seem to result in retention of the mutant precursor in the ER. In the study by Ito *et al.* (32), however, NP II and ER protein were stained separately, and colocalization was judged by the localization of fluorescent material in the cell.

Although the results of this *in vitro* study help to clarify

THIS PAGE BLANK (USPTO)

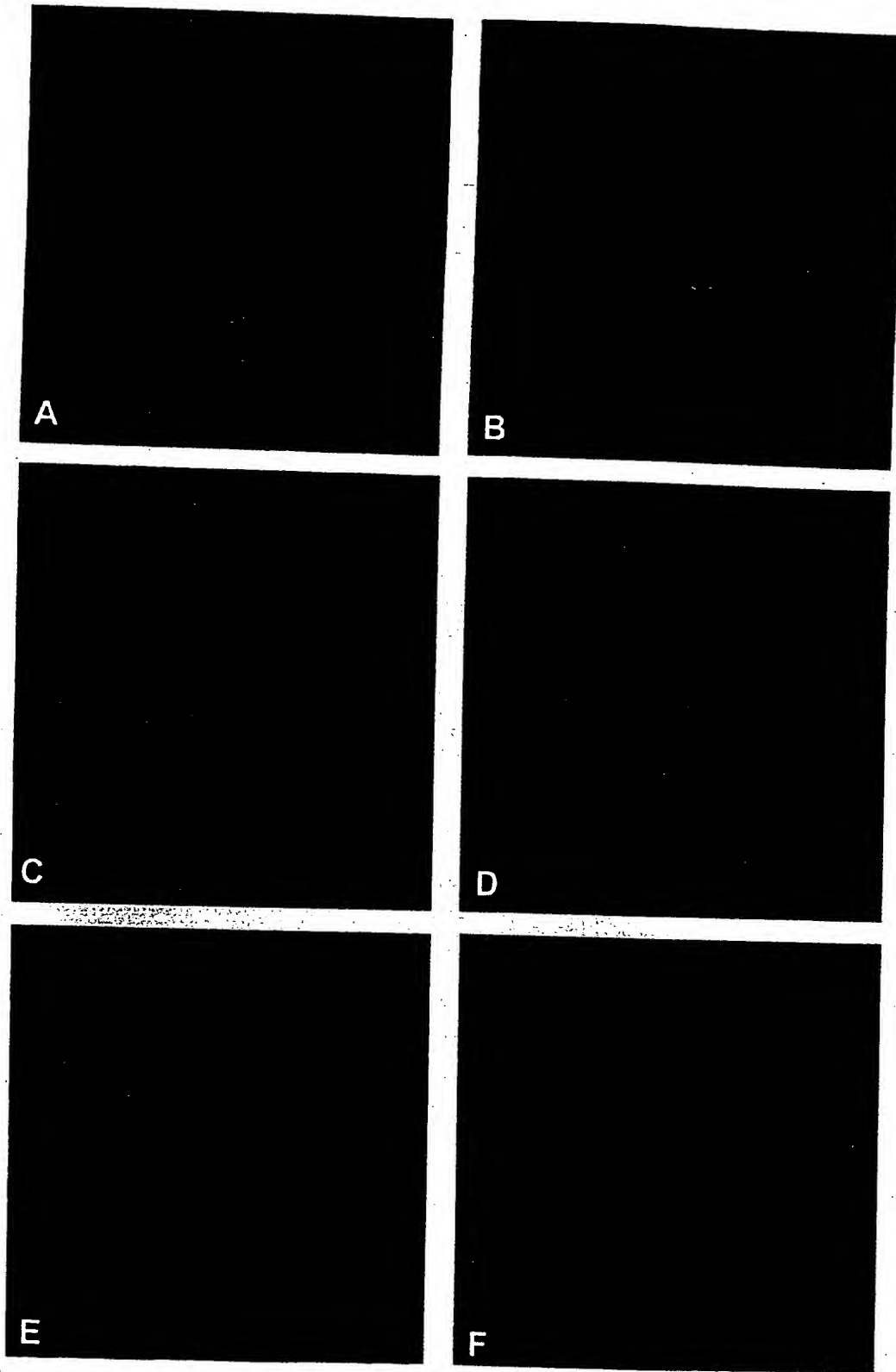


FIG. 7. Confocal laser scanning microscopy analysis of immunostained neuro-2A cells transiently transfected with either wild-type (A, C, and E) or mutant (B, D, and F) cDNA. The cells were double labeled with NPII antibody and an ER protein (Grp78) antibody. These antibodies were visualized by two different fluorescence-conjugated secondary antibodies, seen as *green* (NPII, A and B) and *red* (Grp78, C and D), respectively. The cells expressing the normal precursor were stained throughout the cytoplasm with small aggregates of NPII immunoreactivity, especially in the tips of the cellular processes where the secretory vesicles are normally located (A). However, cells expressing the mutant precursor localized to the perinuclear network and was observed to a much lesser extent in the tips of the cells (B). Colocalization of the two antibodies is visualized as a merge of the two colors to *orange*, which was seen in the mutant cells (F), but not in the wild-type cells (E).

THIS PAGE BLANK (USPTO)

some of the cellular effects of this mutation, they do not demonstrate the mechanism of its dominant negative effect *in vivo*. Autopsy studies in a few patients with adFNDI have consistently shown atrophy of the posterior pituitary, gliosis, and a paucity of magnocellular neurons in the supraoptic nucleus (7–10). These observations as well as the delayed onset of the AVP deficiency (6) suggest that the disease is due to postnatal degeneration of vasopressin-producing neurons in the neurohypophysis. This degeneration might occur because the mutant precursor forms misfolded aggregates that accumulate in the ER and interfere with processing of other proteins essential for cell survival. Preliminary evidence for such a toxic effect has been obtained from other *in vitro* expression studies (32, 48). However, it is also possible that degeneration of neurosecretory neurons is an inconstant, late, or incidental event and that AVP production by the normal allele is impaired primarily because normal and mutant precursors interact to form heterodimers that also cannot be folded or processed properly and are, therefore, rapidly degraded. Moreover, neither mechanism for the dominant negative effect is completely sufficient to explain why the capacity to produce AVP appears to be normal for the first few years of life (6). Further investigation of these patients as well as of *in vitro* cell systems in which both normal and mutant cDNA are stably coexpressed may help to clarify these issues.

Acknowledgments

The authors thank Dorte Rønde, Jane Hagelskjær Knudsen, and Kirsten Tønder for skilled laboratory assistance.

References

- Baylis PH, Robertson GL. 1981 Vasopressin function in familial cranial diabetes insipidus. *Postgrad Med J*. 57:36–40.
- Kaplowitz PB, D'Ercole AJ, Robertson GL. 1982 Radioimmunoassay of vasopressin in familial central diabetes insipidus. *J Pediatr*. 100:76–81.
- Blackett FR, Seif SM, Altmiller DH, Robinson AG. 1983 Familial central diabetes insipidus: vasopressin and nicotine stimulated neurophysin deficiency with subnormal oxytocin and estrogen stimulated neurophysin. *Am J Med Sci*. 286:42–46.
- Pedersen EB, Lamm LU, Albertsen K, et al. 1985 Familial cranial diabetes insipidus: A report of five families. Genetic, diagnostic and therapeutic aspects. *Q J Med*. 57:883–896.
- Os I, Aakesson L, Enger E. 1985 Plasma vasopressin in hereditary cranial diabetes insipidus. *Acta Med Scand*. 217:429–434.
- McLeod JF, Kovacs L, Gaskill MB, Rittig S, Bradley GS, Robertson GL. 1993 Familial neurohypophyseal diabetes insipidus associated with a signal peptide mutation. *J Clin Endocrinol Metab*. 77:599A–599G.
- Forssman H. 1945 On hereditary diabetes insipidus with special regard to a sex-linked form. *Acta Med Scand*. 159:1–196.
- Braverman LE, Mancini JP, McGoldrick DM. 1965 Hereditary diabetes insipidus: a case report with autopsy findings. *Ann Intern Med*. 63:503–508.
- Green J, Buchan G, Alvord JE, Swanson A. 1997 Hereditary and idiopathic types of diabetes insipidus. *Brain*. 90:707–714.
- Bergeron C, Kovacs L, Ezrin C, Mizzen C. 1991 Hereditary diabetes insipidus: an immunohistochemical study of the hypothalamus and pituitary gland. *Acta Neuropathol Berl*. 81:345–348.
- Ito M, Mori Y, Oiso Y, Saito H. 1991 A single base substitution in the coding region for neurophysin II associated with familial central diabetes insipidus. *J Clin Invest*. 87:725–728.
- Bahnsen U, Oosting P, Svaab DF, Nahke P, Richter D, Schmale H. 1992 A missense mutation in the vasopressin-neurophysin precursor gene cosegregates with human autosomal dominant neurohypophyseal diabetes insipidus. *EMBO J*. 11:19–23.
- Ito M, Oiso Y, Murase T, et al. 1993 Possible involvement of inefficient cleavage of preprovasopressin by signal peptidase as a cause for familial central diabetes insipidus. *J Clin Invest*. 91:2565–2571.
- Yuasa H, Ito M, Nagasaki H, et al. 1993 Glu-47, which forms a salt bridge between neurophysin-II and arginine vasopressin, is deleted in patients with familial central diabetes insipidus. *J Clin Endocrinol Metab*. 77:600–604.
- Repaske DR, Browning JE. 1994 A *de novo* mutation in the coding sequence for neurophysin-II (Pro²⁴→Leu) is associated with onset and transmission of autosomal dominant neurohypophyseal diabetes insipidus. *J Clin Endocrinol Metab*. 79:421–427.
- Nagasaki H, Ito M, Yuasa H, et al. 1995 Two novel mutations in the coding region for neurophysin-II associated with familial central diabetes insipidus. *J Clin Endocrinol Metab*. 80:1352–1356.
- Rittig S, Robertson GL, Siggaard C, et al. 1996 Identification of 13 new mutations in the vasopressin-neurophysin II gene in 17 kindreds with familial autosomal dominant neurohypophyseal diabetes insipidus. *Am J Hum Genet*. 58:107–117.
- Rauch F, Lenzner C, Nurnberg P, Frommel C, Vetter U. 1996 A novel mutation in the coding region for neurophysin-II is associated with autosomal dominant neurohypophyseal diabetes insipidus. *Clin Endocrinol (Oxf)*. 44:45–51.
- Ueta Y, Taniguchi S, Yoshida A, et al. 1996 A new type of familial central diabetes insipidus caused by a single base substitution in the neurophysin II coding region of the vasopressin gene. *J Clin Endocrinol Metab*. 81:1787–1790.
- Repaske DR, Summar ML, Krishnamani MR, et al. 1996 Recurrent mutations in the vasopressin-neurophysin II gene cause autosomal dominant neurohypophyseal diabetes insipidus. *J Clin Endocrinol Metab*. 81:2328–2334.
- Rutishauser J, Boni SM, Boni J, et al. 1996 A novel point mutation in the translation initiation codon of the pre-pro-vasopressin-neurophysin II gene cosegregation with morphological abnormalities and clinical symptoms in autosomal dominant neurohypophyseal diabetes insipidus. *J Clin Endocrinol Metab*. 81:192–198.
- Repaske DR, Medlej R, Gultekin EK, et al. 1997 Heterogeneity in clinical manifestation of autosomal dominant neurohypophyseal diabetes insipidus caused by a mutation encoding Ala-1→Val in the signal peptide of the arginine vasopressin/neurophysin II/copeptin precursor. *J Clin Endocrinol Metab*. 82:51–56.
- Gagliardi PC, Bernasconi S, Repaske DR. 1997 Autosomal dominant neurohypophyseal diabetes insipidus associated with a missense mutation encoding Gly²³→Val in neurophysin II. *J Clin Endocrinol Metab*. 82:3643–3646.
- Calvo B, Bilbao JR, Urrutia I, Eizaguirre J, Gaztambide S, Castano L. 1998 Identification of a novel nonsense mutation and a missense substitution in the vasopressin-neurophysin II gene in two Spanish kindreds with familial neurohypophyseal diabetes insipidus. *J Clin Endocrinol Metab*. 83:995–997.
- Heppner C, Kotzka J, Bullmann C, Krone W, Muller WD. 1998 Identification of mutations of the arginine vasopressin-neurophysin II gene in two kindreds with familial central diabetes insipidus. *J Clin Endocrinol Metab*. 83:693–696.
- Rittig S, Siggaard C, Ozata M, et al. 1997 Substitution of Tyr-2 of the anti-diuretic hormone is associated with familial neurohypophyseal diabetes insipidus [Abstract]. *J Invest Med*. 44:387A.
- Hansen L, Rittig S, Robertson GL. 1997 Genetic basis of familial neurohypophyseal diabetes insipidus. *Trends Endocrinol Metab*. 8:363–372.
- Goking NQ, Chertow B, Robertson GL, Rittig S, Siggaard C, Pedersen EB. 1997 Familial neurohypophyseal diabetes insipidus: a novel mutation presenting with enuresis [Abstract]. *J Invest Med*. 45:29A.
- Mundshenk J, Rittig S, Hensen J, Lehnert H. 1998 A novel mutation in a coding region of the arginine vasopressin-neurophysin II (AVP-NP-II) gene in a family with autosomal dominant neurohypophyseal diabetes insipidus [Abstract]. *Exp Clin Endocrinol Diabetes*. 106(Suppl):81.
- Grant FD, Ahmadi A, Hosley CM, Majzoub JA. 1998 Two novel mutations of the vasopressin gene associated with familial diabetes insipidus and identification of an asymptomatic carrier infant. *J Clin Endocrinol Metab*. 83:3958–3964.
- Sausville E, Carney D, Battey J. 1985 The human vasopressin gene is linked to the oxytocin gene and is selectively expressed in a cultured lung cancer cell line. *J Biol Chem*. 260:10236–10241.
- Ito M, Jameson JL. 1997 Molecular basis of autosomal dominant neurohypophyseal diabetes insipidus. Cellular toxicity caused by the accumulation of mutant vasopressin precursors within the endoplasmic reticulum. *J Clin Invest*. 99:1897–1905.
- Olias G, Richter D, Schmale H. 1996 Heterologous expression of human vasopressin-neurophysin precursors in a pituitary cell line: defective transport of a mutant protein from patients with familial diabetes insipidus. *DNA Cell Biol*. 15:929–935.
- Zerbe RL, Robertson GL. 1981 A comparison of plasma vasopressin measurements with a standard indirect test in the differential diagnosis of polyuria. *N Engl J Med*. 305:1539–1546.
- Pedersen EB, Danielsen H, Spencer ES. 1984 Effect of indapamide on renal plasma flow, glomerular filtration rate, and arginine vasopressin in plasma in essential hypertension. *Eur J Clin Pharmacol*. 26:543–547.
- Mohr E, Hillers M, Ivell R, Haulica I, Richter D. 1985 Expression of the vasopressin and oxytocin genes in human hypothalamus. *FEBS Lett*. 193:12–16.
- Kuipers OP, Boot HJ, Vos WM. 1991 Improved site-directed mutagenesis method using PCR. *Nucleic Acids Res*. 19:4558.
- Graham FM, Eb AJ. 1973 A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology*. 52:456–467.
- Jensen TG, Andresen BS, Bross P, et al. 1992 Expression of wild-type and mutant medium-chain acyl-CoA dehydrogenase (MCAD) cDNA in eucaryotic cells. *Biochim Biophys Acta*. 1180:65–72.

THIS PAGE BLANK (USPTO)

40. Jensen TG, Andresen BS, Jensen HK, et al. 1996 Rapid characterization of disease-causing mutations in the low density lipoprotein receptor (LDL-R) gene by overexpression in COS cells. *Ger J Gastroenterol.* 34:1311-1325.
41. Harlow E, Lane D. 1988 Antibodies: a laboratory manual. Cold Spring Harbor: Cold Spring Harbor Laboratory; 418.
42. Munro S, Pelham HR. 1987 A C-terminal signal prevents secretion of luminal ER proteins. *Cell.* 48:899-907.
43. Ledingham JG, Crowe MJ, Forsling ML, Phillips PA, Rolls BJ. 1987 Effects of aging on vasopressin secretion, water excretion, and thirst in man. *Kidney Int.* 21(Suppl):S90-S92.
44. Robertson GL. 1976 The regulation of vasopressin function in health and disease. *Recent Prog Horm Res.* 33:333-385.
45. de Bree FM, Burbach JP. 1994 Heterologous biosynthesis and processing of preprovasopressin in Neuro2A neuroblastoma cells. *Biochimie.* 76:315-319.
46. de Bree FM, van der Kleij AAM, Burbach JPH. 1995 Functional role of hormone-neurophysin association in hormone biosynthesis. In: Saito T, Kurokawa K, Yoshida S, eds. *Neurohypophysis: recent progress of vasopressin and oxytocin research.* Amsterdam: Elsevier; 143-148.
47. Robertson GL, Rittig S, Gu W, et al. 1995 Pathogenesis and pathophysiology of familial neurohypophyseal diabetes insipidus. In: Saito T, Kurokawa K, Yoshida S, eds. *Neurohypophysis: recent progress of vasopressin and oxytocin research.* Amsterdam: Elsevier; 593-603.
48. Robertson GL, Gu W, Gaskill MB, et al. 1996 A vasopressin-neurophysin gene (VPNP) mutation linked to familial neurohypophyseal diabetes insipidus (FNDI) impairs the maturation and vasopressin (VP) production of stably transfected neuroblastoma 2A (N2A) cells [Abstract]. *J Invest Med.* 44:266A.
49. Fikes JD, Barkocy-Gallagher GA, Klapper DG, Bassford Jr PJ. 1997 Maturation of *Escherichia coli* maltose-binding protein by signal peptidase I *in vivo*. *J Biol Chem.* 265:3417-3423.
50. Laforet GA, Kendall DA. 1991 Functional limits of conformation, hydrophobicity, and steric constraints in prokaryotic signal peptide cleavage regions. *J Biol Chem.* 266:1326-1334.
51. Breslow E. 1993 The conformation and functional domains of neurophysins. In: Gross P, Richter D, Robertson GL, eds. *Vasopressin.* Paris: Libbey Eurotext; 143-157.

THIS PAGE BLANK (USPTO)

CONTENTS

Study on Genetic Epidemiology of Albinism.....	Gong Yaoqin, Shao Changshun, Zheng Hong, Chen Bingxi and Guo Yishou (172)
Study on Polymorphism of Ribosomal RNA Gene in Chinese Hans and Yis	Yuan Saying, Jiang Weihong, Lu Yan, Fei Hongming and Chen Renbiao (178)
Invasion of <i>Drosophila albomicans</i> into Shanghai and Areas Nearby and a Study on its Mitochondrial DNA Polymorphism	Chen Weijing, Zhang Jianzhi, Geng Zhencheng and Zhu Dingliang (187)
Chromosome Abnormality and Dominant Lethality among the F ₁ Generation of Mice after γ -ray Irradiation	Yao Suyan, Zhang Chaoyang, Dai Lianlian, Gao Changwen and Wei Lixin (192)
Estimation of Evolutionary Distances Between Protein Sequences	Yang Ziheng (200)
A Cytological Observation on the Pollen Abortion in the Bilinear Nuclear Sterile Strain of the New-type Uplandrice Kunzhi S-1.....	Wu Shibin, Sun Youming, Zhou Kaiyuan and Wan Jianhui (204)
Construction of a Rice Molecular Linkage Map Using a Double Hploid (DH) Population	Xu Jichen, Zhu Lihuang, Chen Ying, Lu Chaofu and Cai Hongwei (214)
Improvement of Plant Regeneration frequency <i>in vitro</i> In Indica Rice.....	Tian Wenzhong, Iann Rance, Elunialai Sivamani, Claude Fauquet and Roger N. Beachy (221)
Effect of Genomes on the Mesophyll Protoplast Culture and Plant Regene- ration of <i>Brassica</i>	Li Shijun, Meng Zheng and Li Debao (226)
Effect of the Changes of Amino Acids on Both Signal Peptide C-Terminal and Mature Protein N-Terminal Region to the Secretion of α -Amylase in <i>B. subtilis</i>	Ji Yonggang, Yang Lizhu, Chen Qimin, Ma Ming, Geng Yunqi and Jiang Ruzhang (234)
Construction of Secretion Vectors Using Signal Sequence of <i>Bacillus</i> <i>subtilis</i> Alkaline Protease E Gene	Liu Yongliang and Tong Kezhong (246)

THIS PAGE BLANK (USPTO)

信号肽 C-端和成熟蛋白 N-端氨基酸 序列的变化对 α -淀粉酶分泌作用的影响^①

纪永刚 杨丽珠 陈启民

马明 耿运琪 蒋如璋

(南开大学生物工程研究室 天津 300071)

摘要 本文利用 pAmy 413 质粒通过定点诱变技术, 在 α -淀粉酶基因的 287 和 291 位分别引入 1 个 G 和 C, 代替原来的 T 和 G, 构建成质粒 pAmy413B, 使成熟的 α -淀粉酶 N-端 'Leu'Met 变为 'Arg'Ile。pAmy413L α -淀粉酶信号序列因引入 1 个多酶切点接头而比 pAmy 413 的 29 个氨基酸增加了 13 个氨基酸, 创造了 1 个新的信号肽酶 I 识别窗口 Ala-Gln-Ala⁺Ser。利用计算机对该信号肽进行了二级结构分析。这两株突变株产酶能力分别为野生株 (pAmy413) 的 3% 和 36%; 胞外 α -淀粉酶的分子量同于野生株; 末端分析显示, 成熟酶第一个氨基酸为 Ala, 表明信号肽酶 I 在野生型识别窗口 Ala-Ala-Ala⁺Ala 处切割。结果还表明, *B. licheniformis* α -淀粉酶在 *B. subtilis* 中分泌作用属共翻译转移模型。

关键词 信号肽, α -淀粉酶, 定点诱变, 蛋白质分泌, 共翻译转移

蛋白质分泌是极其复杂的生化反应过程。近年来分子遗传学、生物化学、生物膜等方面的研究虽然取得很大进展, 但对蛋白质分泌的分子机制仍没有完整的认识。关于原核生物蛋白质分泌机制的知识主要来自 *E. coli*。芽孢杆菌具有高活性蛋白质分泌系统, 而蛋白质分泌机制的研究还刚刚开始。主要研究集中在信号肽的结构和功能^[6,10], 构建分泌型载体^[2,10,11,17]和构建多重蛋白酶缺陷菌株^[18]。

本实验室已经克隆了 *B. licheniformis* α -淀粉酶基因^[9], 其核苷酸序列也已清楚^[19]。由核苷酸序列推知信号肽由 29 个氨基酸残基组成。信号肽酶的切割窗口为 Ala-Ala-Ala⁺Ala。本文以此为材料, 采用重组 DNA 技术和定点诱变, 研究了信号肽 C-端和成熟蛋白 N-端氨基酸序列的变化对 α -淀粉酶分泌作用的影响。

1 材料和方法

1.1 菌株、质粒、噬菌体和培养基

本研究所用菌株、质粒和噬菌体见表 1。LB 培养基^[13]用于细菌培养及淀粉酶活性测定。含 0.5% 可溶性淀粉的 LB 固体培养基用于 α -淀粉酶活性平板检测。SMMP 和 DM3 培养基^[7]用于 *B. subtilis* 原生质体转化。2 × YT 培养基^[12]用于噬菌体扩增。

1.2 酶及试剂

限制酶、T4DNA 连接酶、T4DNA 多核苷酸激酶、Klenow 酶、Mung Bean 核酸酶分别购自 New England Biolabs 和中国医学科学院基础所。IPTG 和 X-gal 购自

① 本文于 1992 年 5 月 15 日收到。本文是国家自然科学基金资助

THIS PAGE BLANK (USPTO)

表 1 菌株、质粒和噬菌体
Table 1 Strain, plasmid and phage

菌株、质粒和噬菌体 Strain, plasmid and phage	基因型或表现型 Genotype or Phenotype	来源 Source
菌株 Strains <i>E. coli</i> JM109	recA1 Δ lacpro endA1 gyrA96 thi-1 hsdR17 supE44 relA1 F' traD36 proAB ⁺ lacI ^q Z Δ M15	本室保藏 Lab. stock
<i>E. coli</i> RE1032	dut ung F'	北京大学生物系 Dept. of Biology Beijing University
<i>B. subtilis</i> AS1.1176	str trp amy	中科院微生物研究所 Institute of Microbiology Academic Sinica
质粒 Plasmids pAmy413 pAmy413B pAmy413L	Km ^r Amy ⁺ Km ^r Amy ⁺ Km ^r Amy ⁺	本室构建 This lab. 本文构建 This study 本室构建 This Lab.
噬菌体 Phage M ₁₃ mp18		本室保藏 Lab. stock

Sigma。dNTP 及 ATP 购自 Pharmacia。DNS-Cl 购自 Fluka。可溶性淀粉购自 BDH。其它药品均为国产试剂。

1.3 DNA 制备与纯化

M₁₃mp18 RF DNA 和 *B. subtilis* 质粒 DNA 提取按 Sambrook 法^[13]。DNA 片段经琼脂糖凝胶电泳后,用 Bio-Rad 公司 422 型电洗脱仪制备,并用 Beckman Du-7 型紫外分光光度计测定 DNA 浓度。

1.4 定点诱变

合成 25 聚体引物 5'-AATGGGACGCGGATCCAGTATTTTG-3', 并按 Sambrook 法^[13]磷酸化;培养含尿嘧啶重组噬菌体,并按 ABI 公司电传资料制备其单链 DNA。采用缺口双链法^[14]完成定点诱变。

1.5 遗传转导与转化

E. coli 遗传转导按 Sambrook 法^[13]。*B. subtilis* 原生质体转化按 Chang 法^[7]。用 DM3 培养基筛选抗性转化子, Km 浓度为 1mg/ml。

1.6 淀粉酶阳性菌的检出和活性测定

采用碘蒸汽法^[15]显示淀粉酶阳性菌落。以二硝基水杨酸法^[13]测定 α -淀粉酶活性。85℃ 条件下,每毫升酶液每分钟分解出 1 微摩尔还原糖定义为 1 个酶活单位。

1.7 α -淀粉酶分部提取和纯化

按周毅法^[3]分部提取 α -淀粉酶。发酵液分级盐析,然后经 DEAE 和 CM 纤维素柱



THIS PAGE BLANK (USPTO)

层析纯化^[4]。

1.8 分子量测定

SDS-聚丙烯酰胺凝胶电泳按 Laemmli 法^[5], 采用 10% 连续胶, 用考马斯亮蓝染色显带。

1.9 蛋白质 N-末端分析

采用 DNS-Cl 法^[6]制备 DNS-氨基酸, 以聚酰胺薄膜进行双向层析。展层液 I: 1.5% 甲酸水溶液; II: 苯-醋酸 (9:1, V/V)。在 254nm 紫外光下, DNS-氨基酸显示黄色荧光斑点。

1.10 多肽链二级结构分析

采用 IBM/PCXT 计算机及 MICROGENIE 软盘的程序分析。

2 结果

2.1 pAmy413B 质粒的构建

为了研究 α -淀粉酶成熟蛋白 N-端氨基酸替换对蛋白质分泌的影响, 我们设计了一个 25 聚体引物, 通过定点诱变, 在 pAmy413 α -淀粉酶基因的第 287 和 291 位分别引入 G 和 C 代替原来的 T 和 G, 使 α -淀粉酶成熟蛋白的 ¹Leu⁸Met 变成 ¹Arg⁸Ile, 同时在 287—292 位创造了 BamHI 切点。这样使成熟蛋白的前 15 个氨基酸残基的电中性变为荷 1 个正电荷, 用 BamHI 进行酶切分析的结果证实了预期的核苷酸改变。这个质粒定名为 pAmy413B, 过程如图 1。

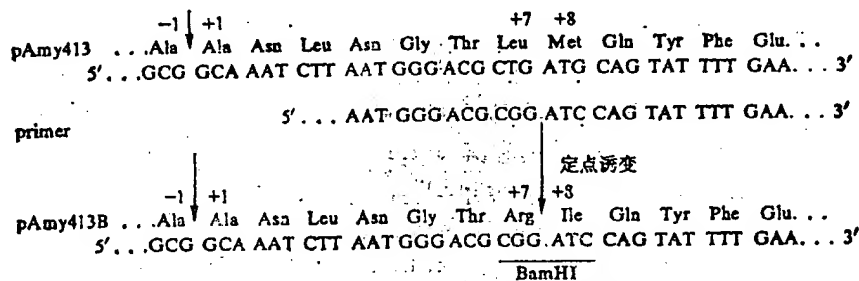


图1 质粒 pAmy413B 的构建

Fig. 1 Construction of plasmid pAmy413B

↓ 示信号肽酶 I 识别切割位点

↓ Shows the recognition and cleavage site of signal peptidase I

2.2 pAmy413L α -淀粉酶的信号序列及二级结构分析

为了研究 α -淀粉酶信号肽的功能, 在 pAmy413 α -淀粉酶基因的信号序列区的 PstI 切点引入 pUC13 的多酶切点接头。经过读框修改, 构建成质粒 pAmy413L。pAmy413L α -淀粉酶的信号序列和二级结构见图 2。

从图 2 可见, α -淀粉酶的信号肽由原来的 29 个氨基酸增加了 13 个氨基酸, 同时在一13 到一16 位引进了新的信号肽酶 I 的识别窗口, 即 Ala-Gln-Ala⁴Ser。二级结构分析显示, 在新增加的 13 个氨基酸残基的中部 (一7—一10 位) 存在转角结构。这在一般信号肽中少见。

THIS PAGE BLANK (USPTO)

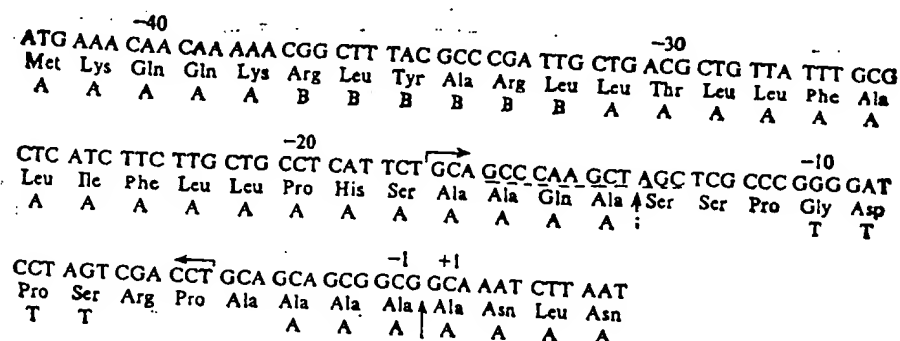
图2 pAmy413L α -淀粉酶的信号序列及二级结构

Fig. 2 The signal sequence and its secondary structure of plasmid pAmy413L α -amylase
 ↑ 切割位点; ↑ 人工构建的切割位点; [X] 多酶切点接头; A: α -螺旋; B: β -折叠; T: β -转角。
 ↑ cleavage site; ↑ artificial cleavage site, [X] polylinker; A: α -helix; B: β -sheet; T: β -turn.

2.3 淀粉酶基因的表达

为了探查信号肽序列的改变和 α -淀粉酶成熟蛋白 N-端氨基酸替换对 α -淀粉酶分泌的影响,进行了平板检测和定位分析。

2.3.1 平板检测 将 3 个等基因株 *B. subtilis* (pAmy413), (pAmy413B), (pAmy413L) 点接在含 0.5% 淀粉的 LB 平板上, 37°C 培养过夜。以碘蒸汽显示淀粉水解圈 (图 3)。从水解圈显示 α -淀粉酶的分泌量为 pAmy413 和 pAmy413L 都大于 pAmy413B。

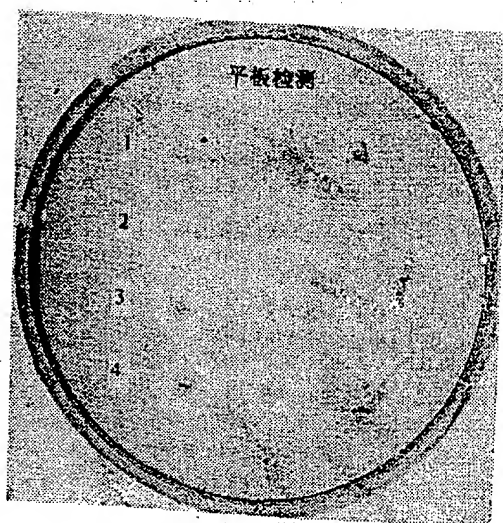


图3 不同等基因菌株 α -淀粉酶的分泌能力
 Fig. 3 α -amylase secretion ability in different isogenic strains

1. *B. subtilis* AS 1.1176 (control) 2. *B. subtilis* (pAmy413B) 3. *B. subtilis* (pAmy413L) 4. *B. subtilis* (pAmy413)

2.3.2 α -淀粉酶定位 将上述 3 个菌株分别接种于 LB 培养液, 按材料和方法定时测定胞外及总酶活, 结果从图 4、5、6 可以看出, pAmy413L, pAmy413B 酶活分别为 1U/ml 和 0.07 U/ml, 相当于野生型 pAmy413 (2.8U/ml) 的 36% 和 3%。并且产生的 α -淀粉酶全部分泌到胞外。

2.4 α -淀粉酶分子量比较

将 pAmy413、pAmy413L 和 pAmy413B 的发酵液, 按材料和方法纯化 α -淀粉酶, 聚丙烯酰胺凝胶电泳发现成熟的 α -淀粉酶分子量均为 55000D (图 7)。说明信号肽酶 I 对 pAmy413, pAmy413L, pAmy413B 产生的原淀粉酶

均以同样的方式和位点进行加工, 产生等分子量的成熟的 α -淀粉酶。

2.5 成熟 α -淀粉酶末端分析

将上述 3 菌株所产 α -淀粉酶纯化后, 以 DNS-Cl 法分别进行 N-末端氨基酸分析, 结果如图 8。结果表明 3 株菌所产生的 α -淀粉酶 N-端均为 Ala。说明信号肽酶 I 的识别窗口为 Ala-Ala-Ala⁺Ala 这一结果与分子量测定结果相符。但是并未检出 pAmy413L

THIS PAGE BLANK (USPTO)

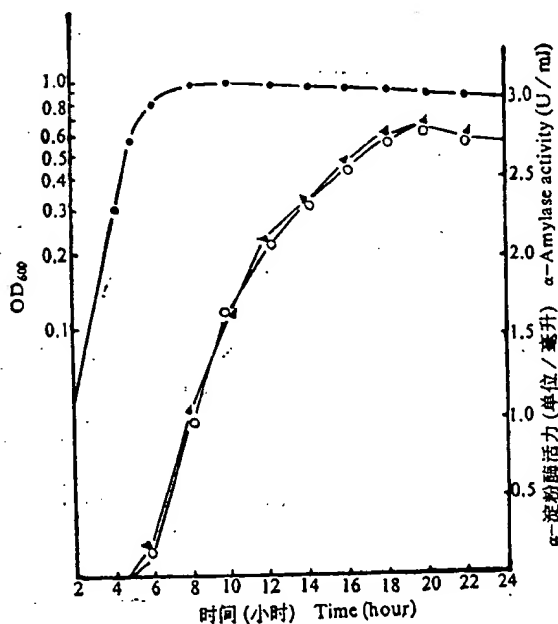


图4 *B. subtilis*(pAmy413) 胞外和总酶活增长曲线
Fig. 4 Growth and enzyme activity curves of *B. subtilis* (pAmy413)

—●—●—生长曲线; —▲—▲—总酶活; —○—○—胞外酶活
—●—●—Growth curve; —▲—▲—Total activity; —○—○—Extracellular activity

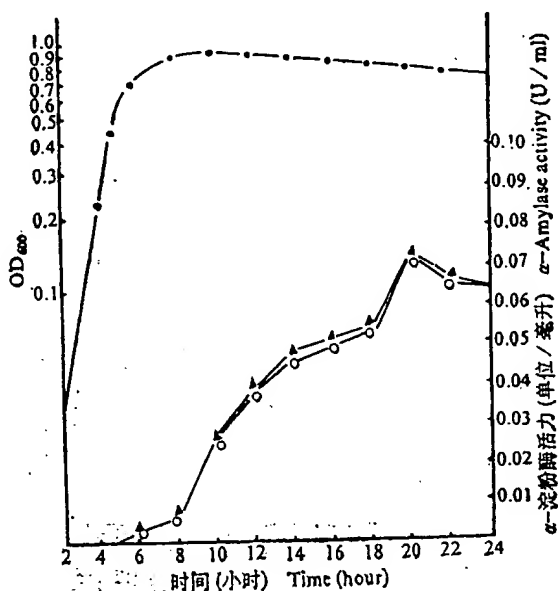


图5 *B. subtilis* (pAmy413B) 胞外和总酶活增长曲线
Fig. 5 Growth and enzyme activity curves of *B. subtilis* (pAmy413B)

—●—●—生长曲线; —▲—▲—总酶活; —○—○—胞外酶活
—●—●—Growth curve; —▲—▲—Total activity; —○—○—Extracellular activity

成熟淀粉酶 N-端的 Ser, 说明 pAmy413L 产生的原 α -淀粉酶被信号肽酶 I 最终识别窗口是 Ala-Ala-Ala⁺Ala, 而不是 Ala-Gln-Ala⁺Ser。

3 讨论

在 pAmy413L 的 α -淀粉酶信号序列中含有两个信号肽酶 I 识别窗口, 人工创造的 Ala-Gln-Ala⁺Ser 和野生型的 Ala-Ala-Ala⁺Ala。但成熟酶 N-端氨基酸为 Ala。说明在人工创造的识别窗口存在下, 信号肽酶优先选择野生型的识别窗口, 致使 3 菌株的成熟 α -淀粉酶 N-端均为 Ala; 或许信号肽酶在两个窗口同时迅速切割, 无法分别检出两种成熟蛋白的 N-末端氨基酸。多酶切点片段的引入, 也有可能致信号肽酶 I 的原加工位点由 Ala-Ala-Ala-Ala⁺Ala 变成 Ala-Ala-Ala⁺Ala-Ala。这时, 采用 PAGE 电泳或单个氨基酸的末端分析就不足以确定成熟蛋白质间的这种微小差别。而应对成熟蛋白 N-端多个氨基酸进行分析, 方能显示信号肽酶 I 的识别切割位点是否有这种改变。这方面工作正在进行中。

pAmy413L α -淀粉酶信号肽 C-端增加了 13 个氨基酸残基, 并创造了信号肽酶 I 的识别序列 Ala-Gln-Ala⁺Ser。二级结构分析表明: 在新增加的 13 个氨基酸中部(-7—10 位), 存在转角结构, 这种结构在一般信号肽中少见。同时在一6 和一9 位分别有一荷正电的 Arg 和荷负电的 Asp, 这仍使信号肽 C-端的净电荷为零。这样加长的信号肽致使 α -淀粉酶的产量降为野生型的 36%。这可以解释为: 转角结构的引入, 使信号肽的

THIS PAGE BLANK (USPTO)

Effect of the Changes of Amino Acids on Both
Signal Peptide C-Terminal and Mature
Protein N-Terminal Region to the
Secretion of α -Amylase in
*B. subtilis*①

Ji Yonggang Yang Lizhu Chen Qimin
Ma Ming Geng Yunqi Jiang Ruzhang
(Lab. for Biotechnology, Nankai Univ., Tianjin 300071)

Abstract

By site-directed mutagenesis, G and C have taken the place of T and G at nucleotide sequence 287 and 291 of *B. licheniformis* α -amylase gene to generate pAmy413B and the N-terminal sequence of mature protein have been changed from 'Leu'Met to 'Arg'Ile. By the insertion of polylinker into the C-terminal of the signal sequence of α -amylase gene of pAmy413, the signal peptide of α -amylase produced by pAmy413L is 13 amino acids more than the pAmy413 (which is 29 amino acids long) and also, a new recognition cleavage sequence for signal peptidase I (Ala-Gln-Ala'Ser) is created; The secondary structure of the signal peptide has been analyzed by computer programs. The α -amylase relative activity of the two mutant strains is 3% and 36% of pAmy413, respectively. The molecular weight of extracellular α -amylase is the same as pAmy413. Terminal analysis shows that the N-terminal amino acid of mature protein is Ala, not Ser, and suggests that SPase I prefers to cleavage at the wild type recognition site (Ala-Ala-Ala'Ala). Therefore, all of the above results show that the secretion of α -amylase in *B. subtilis* is in accordance with the co-translational transportation model.

Key words α -amylase, Signal peptide, Site-directed mutagenesis, Protein secretion, Co-translational transfer

① Received May 15, 1992. The Project Supported by National Natural Science Foundation of China

THIS PAGE BLANK (USPTO)

Processing of *Escherichia coli* Alkaline Phosphatase: Role of the Primary Structure of the Signal Peptide Cleavage Region

Andrew L. Karamyshev¹, Zemphyra N. Karamysheva¹,
Andrey V. Kajava², Vladimir N. Ksenzenko³
and Marina A. Nesmeyanova^{1*}

¹Research group "Protein Secretion in Bacteria", Skryabin Institute of Biochemistry and Physiology of Microorganisms Russian Academy of Sciences 142292 Pushchino, Moscow Region, Russia

²Swiss Institute for Experimental Cancer Research Chemin des Boveresses 155 CH-1066, Epalinges s/Lausanne, Switzerland

³Laboratory of Structural and Functional Analysis of Genetic Systems of Microorganisms Skryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia

A wide range (69) of mutant *Escherichia coli* alkaline phosphatases with single amino acid substitutions at positions from –5 to +1 of the signal peptide were obtained for studying protein processing as a function of the primary structure of the cleavage region. Amber suppressor mutagenesis, used to create mutant proteins, included: (i) introduction of amber mutations into respective positions of the *phoA* gene; and (ii) expression of each mutant *phoA* allele in *E. coli* strains producing amber suppressor tRNAs specific to Ala, Cys, Gln, Glu, Gly, His, Leu, Lys, Phe, Pro, Ser and Tyr. Most amino acid substitutions at positions –3 and –1 resulted in a complete block of protein processing. These data give new experimental support for the "–3, –1 rule". Only Ala, Gly and Ser at position –1 allowed protein processing, and Ala provided the highest rate of processing. The results revealed the more conservative nature of the amino acids at the –1 position of signal peptides of Gram-negative bacteria as compared with those of eukaryotic organisms. Position –3 was less regular, since not only Ala, Ser and Gly, but also Leu and Cys at this position, allowed the processing. Mutations at position –4 had an insignificant effect on the processing. Surprisingly, efficient processing was provided mainly by large amino acid residues at position –2 and by middle-sized residues at position –5, indicating that the processing rate is affected by the size of amino acid residues not only at positions –1 and –3. Conformation analysis of the cleavage site taken together with the mutation and statistical data suggests an extended β -conformation of the –5 to –1 region in the signal peptidase binding pocket.

© 1998 Academic Press Limited

*Corresponding author

Keywords: signal peptide; signal peptidase; protein secretion; protein processing; alkaline phosphatase

Introduction

Most bacterial secretory proteins are synthesized in the cytoplasm as protein precursors containing an additional N-terminal sequence called the signal peptide, which plays a crucial role in protein recognition of the cell secretory machinery and in initiation of protein membrane translocation (for review see Pugsley, 1993; Izard & Kendall, 1994).

After translocation, the signal peptide is cleaved by a membrane-bound signal peptidase, and the precursors are converted into mature proteins. It is known that the sites of signal peptide cleavage by leader peptidase are rather regular and described by the –3, –1 rule (von Heijne, 1983) or A-X-B model (Perlman & Halvorson, 1983) originally revealed by statistical evaluation of the known primary structures of signal peptides. In agreement with this rule, only small neutral residues are present at positions –3 and –1. Such structural regularity in the –3, –1 region is considered to be necessary for recognition of the signal peptide cleavage site by leader peptidase but inessential for protein translocation across the cytoplasmic mem-

Present address: A. L. Karamyshev and Z. N. Karamysheva, Department of Tumor Biology, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108, Japan
E-mail: aniram@ibpm.serpukhov.su

THIS PAGE BLANK (USPTO)

brane (Fikes *et al.*, 1990; Pollitt *et al.*, 1986; Shen *et al.*, 1991). In support of this rule it was shown that amino acid substitutions at the above positions of precursors of lipoprotein (Pollitt *et al.*, 1986), phage M13 procoat protein (Kuhn & Wickner, 1985; Shen *et al.*, 1991), and maltose-binding protein (Fikes & Bassford, 1987; Fikes *et al.*, 1990), resulted in prevention of the processing. We have also shown that the substitution of Val for Ala(-1) in the signal peptide of alkaline phosphatase completely inhibits the processing of this mutant protein (Karamyshev *et al.*, 1994b; Nesmeyanova *et al.*, 1994).

At the same time, the effect of some amino acid substitutions at position -3 on the processing may vary from protein to protein. Indeed, mutant maltose-binding proteins with Leu, Thr or Val at position -3 were processed efficiently (Fikes *et al.*, 1990), while the processing of the mutant phage M13 procoat protein with the same substitutions was significantly slower (particularly in the proteins with Leu(-3) or Thr(-3) (Shen *et al.*, 1991). Besides, Pro(-3) of the procoat M13 leads to a complete processing block but exerts no effect on the processing of the maltose-binding protein. Concerning the occurrence of amino acid residues at positions from -6 to -4, Pro and Gly particularly are often found in this boundary region between the hydrophobic core and c-region of the signal peptide (Perlman & Halvorson, 1983; von Heijne, 1986a). However, the role of these residues is still obscure. The frequent occurrence of these residues in the β -turns of water-soluble globular proteins (Chou & Fasman, 1978) resulted in a widely accepted opinion that these residues promote the formation of a β -turn in the signal peptides as well (Rosenblatt *et al.*, 1980; Perlman & Halvorson, 1983). Pro(-6) of the M13 procoat protein was shown to be absolutely essential for the processing, and its replacement by many other amino acid residues resulted in a complete impairment of the signal peptide cleavage (Shen *et al.*, 1991). Mutagenesis studies of the maltose-binding protein showed that a decrease of the processing efficiency in mutant proteins correlated with a low probability of β -turn formation (Barkocy-Gallagher *et al.*, 1994). At the same time, a study of the processing of alkaline phosphatases with a mutated signal peptide lacking proline and containing only helix-fostering residues (Leu or Ala) showed an efficient processing of mutant proteins (Laforet & Kendall, 1991). The authors of the latter work suggest that the hydrophobicity of amino acid residues is more important for processing than the possibility of β -turn formation. However, Shen *et al.* (1991) failed to find a correlation between the processing and hydrophobicity: neither Ala, Gly nor Trp, having hydrophobicity similar to Pro, could be functional substitutes for it at position -6 of the M13 procoat protein (Shen *et al.*, 1991). Moreover, the substitution of Gln, which is more polar than Pro at position -6, slowed down the processing. Thus, the current experimental data are insufficient for com-

plete understanding of the structural basis of signal peptide recognition by signal peptidase: the total size of this site is unknown; it is unclear whether the nature of the amino acid residues is important for the processing at both positions -3, -1 and at nearby positions.

The goal of this work was to study systematically the role of the primary structure in and near the signal peptide cleavage site in the processing of *E. coli* alkaline phosphatase. The objectives were: (1) to obtain a number of mutant proteins with amino acid substitutions at positions -5, -4, -3, -2, -1 and +1 of the alkaline phosphatase signal peptide; (2) to study the effect of the above substitutions on protein processing; (3) to try to refine the interaction of the signal peptide c-region with the binding pocket of signal peptidase, on the basis of new experimental data and using a stereochemical method of analysis.

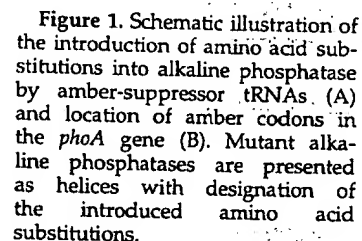
Results

Introduction of amino acid substitutions into the processing region of *E. coli* alkaline phosphatase using amber suppressor tRNAs

Mutant alkaline phosphatases were obtained by: (i) introduction of amber mutations into the alkaline phosphatase (*phoA*) gene and (ii) expression of these mutant *phoA* alleles in *E. coli* strains producing amber suppressor tRNAs that are able to insert an amino acid of interest in response to amber mutation (Figure 1). In these experiments 13 different amber-suppressor tRNAs specific to Ala, Arg, Cys, Gln, Glu, Gly, His, Leu, Lys, Phe, Pro, Ser and Tyr can be used, because, at least in *E. coli*, the other seven amber suppressor tRNAs (Asp, Asn, Ile, Met, Thr, Trp, Val) do not retain their initial amino acid specificity due to substitutions in anticodon sequences (Normanly *et al.*, 1990).

Six amber mutations were independently introduced by site-directed mutagenesis into certain positions of the *phoA* gene corresponding to the codons of the amino acid residues Arg(+1), Ala(-1), Lys(-2), Tyr(-3), Val(-4) and Pro(-5) in the signal peptide cleavage region of the protein. The mutant alleles of *phoA* were cloned in the plasmid vector p15SK(-) which bears a chloramphenicol resistance marker and the p15A *ori* of replication, allowing its use in a two-plasmid system in combination with plasmids having the ColE1 *ori* of replication, which carry cloned genes of amber-suppressor tRNAs. Both natural (with a chromosomal localization) and synthetic (cloned in the plasmid vector pGFIB1 under the control of constitutive *lpp* promoter) amber suppressor tRNA genes were used in this study. Expression of each of the above six mutant *phoA* alleles in 12 *E. coli* amber suppressor strains resulted in 69 different mutant PhoAs (the other three synthesized proteins were the same as wild-type PhoA).

THIS PAGE BLANK (USPTO)



The effect of amino acid substitutions on protein maturation was assessed by the rate of conversion of pulse-labeled mutant protein precursors into the mature form *in vivo* (Figure 2). Amino acid residues at positions -3 and -1 were revealed to be highly conserved (Figures 2 and 3). Most amino acid substitutions at these positions resulted in a processing block. The mutant proteins with Phe(-1), Tyr(-1), Glu(-1), Lys(-1), Leu(-1), Gln(-1) and His(-1) remained as precursors even 60 minutes after [³⁵S]methionine incorporation was stopped. Of those that were processed, the proteins with Cys(-1) and Pro(-1) were processed with the lowest efficiency: only 29% and 11%, of the precursors respectively were converted into the mature form. The least effect on the processing was shown with Gly(-1) and Ser(-1), although the proteins having these amino acids were processed at a lower rate than the wild-type protein. Thus, among 12 amino acid substitutions analyzed Ala(-1) provided the highest rate of processing. Gly(-1) and Ser(-1) also allowed processing but at a lower rate. Position -3 is less conserved. Pro-

Positions +1 and -2 were still less conserved. Only the introduction of Pro(+1) completely inhibited the processing. Pro(-2) allowed processing but at the lowest rate. It would be interesting to note that a certain amount of the protein with Cys(+1) or Cys(-2) remained unprocessed during the period of observation, although the processing of a considerable part of it proceeded rather quickly. Efficient processing was provided mainly by large amino acid residues (Phe, Tyr, Leu, His) at position -2. As to the -4 position, most amino acid substitutions at this position had no effect on the rate of protein maturation. Only Lys(-4) was shown to affect the processing. Position -5, however, appeared to be important for processing. Many amino acids at this position, such as Lys, Tyr, Leu, Ala, His, Phe and Gly, appreciably decreased the processing rate, although none of these amino acids completely blocked the processing. The data show that the primary structure of the cleavage region ranging from the -5 to +1 residues is important for the efficient processing of alkaline phosphatase.

THIS PAGE BLANK (USPTO)

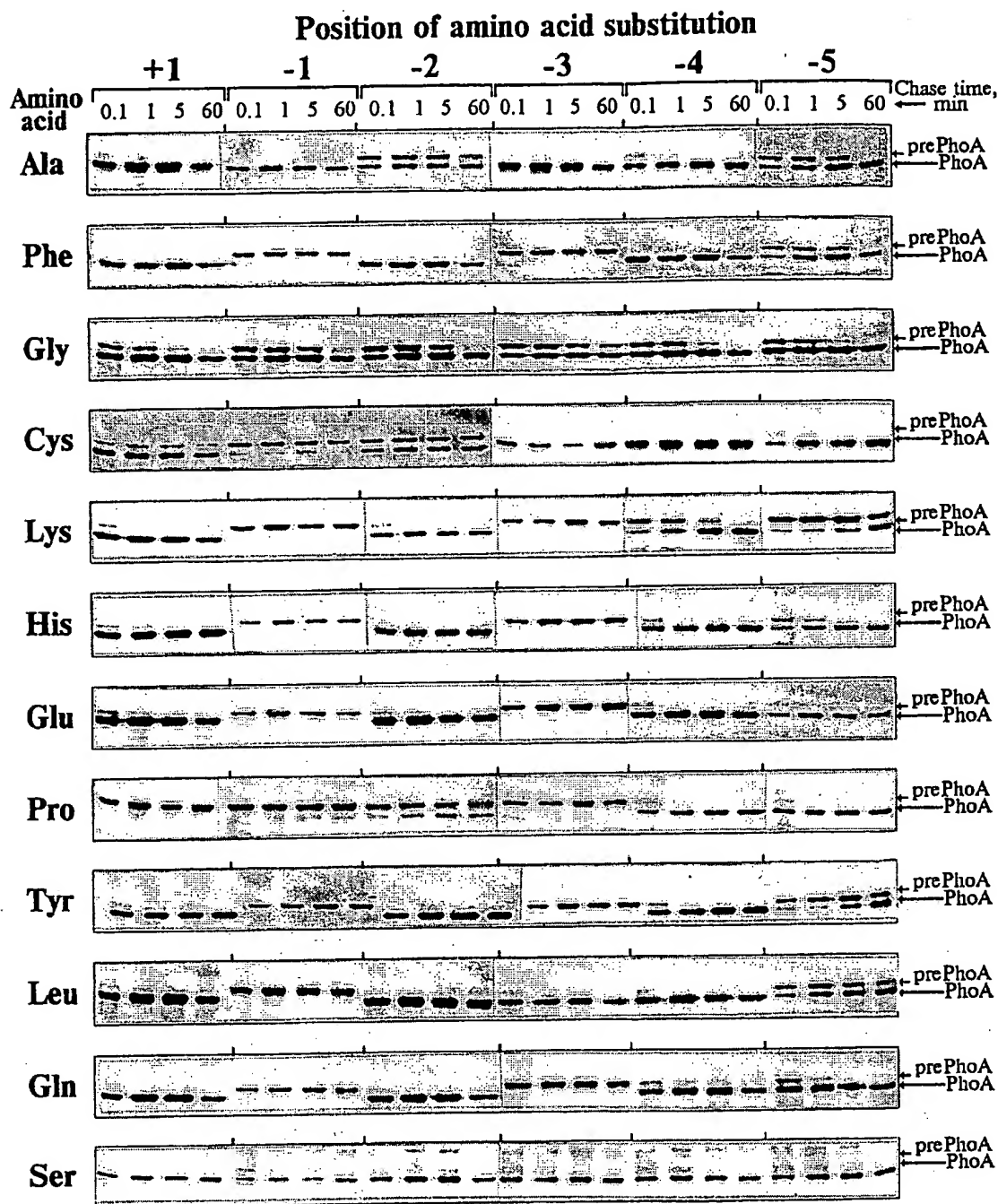


Figure 2. The dynamics of processing of mutant alkaline phosphatases with amino acid substitutions in and near the signal peptide cleavage site. Cells were pulse-labeled with L-[35 S]methionine for 60 seconds, and the radioactivity was chased for indicated periods of time. Pulse-labeled alkaline phosphatase and its precursor were immunoprecipitated using affinity-purified rabbit antibodies against alkaline phosphatase and resolved by 10% SDS-PAGE followed by autoradiography.

Rational for the processing impairment of mutant alkaline phosphatases

The processing impairment may be explained by two possible mechanisms: (1) a block of preprotein

translocation across the membrane, which makes the precursor inaccessible for signal peptidase; and (2) direct inhibition of signal peptide cleavage due to a disturbance of the structure of the protein that acts as a substrate for signal peptidase. Here, the

THIS PAGE BLANK (USPTO)

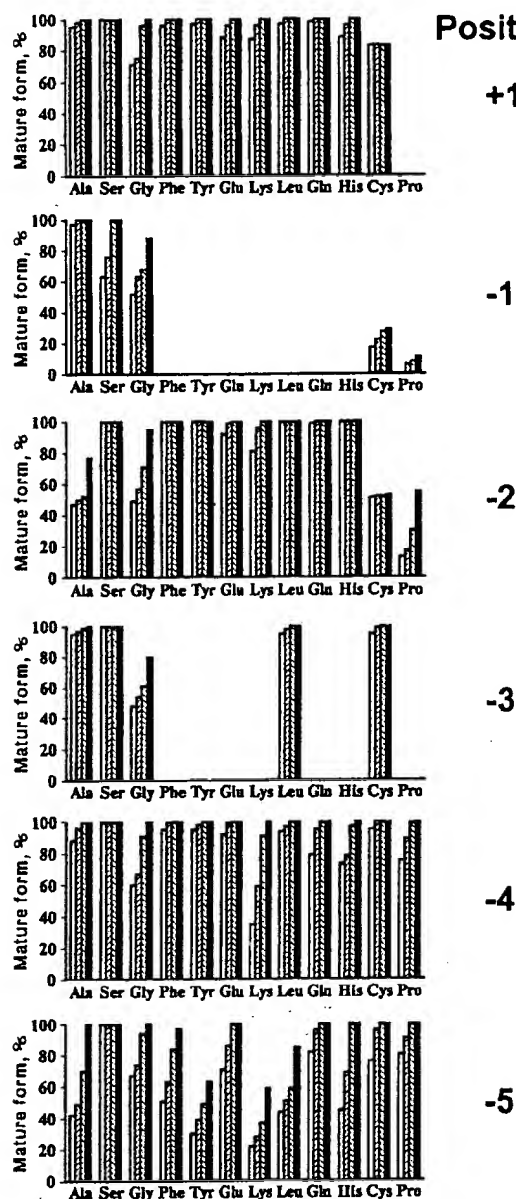


Figure 3. Summary of the quantitative analysis of alkaline phosphatase processing as a function of amino acid substitutions in the signal peptide cleavage region. Relative quantities of mature alkaline phosphatase from the total amount of the mature protein and its precursor are presented. Amino acid substitutions and their positions in the protein are indicated. Proteins with Ala(-1), Lys(-2) and Pro(-5) correspond to wild-type alkaline phosphatase and were obtained using the corresponding amber suppressor. The relative quantity of mature enzyme, with adjustments of additional methionine residue in the precursor, was calculated for 0.1 (open bars), 1.0 (left hatched bars), 5.0 (right hatched bars) and 60.0 (filled bars) minutes of processing time. Quantitation of proteins was performed using an LKB UltroScan laser densitometer. The total amount of the mature alkaline phosphatase and its precursor was set to 100% for each point of the processing time.

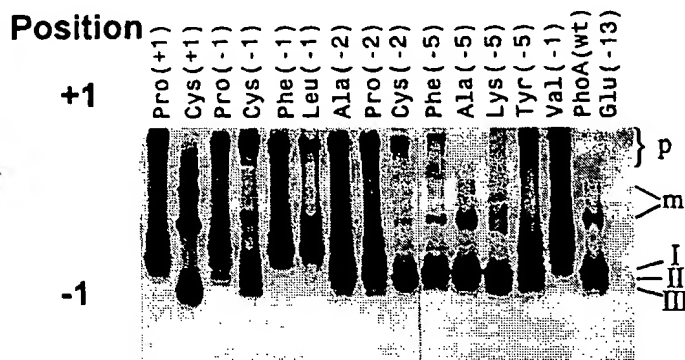


Figure 4. Multiple forms of wild-type and mutant alkaline phosphatases. Samples were analyzed by electrophoresis (7.5% PAGE) under non-denaturing conditions, and the active enzyme was revealed by treatment of the gel with α -naphthyl phosphate as the alkaline phosphatase substrate and Fast Red Dye TR. Isoforms of wild-type alkaline phosphatase (I, II, III), multimer forms (m), and active precursor (p) are indicated.

above mechanisms were differentiated by the analysis of multiple forms of active enzyme. It is known that alkaline phosphatase becomes active only after translocation across the cytoplasmic membrane into the periplasm, where there exist necessary conditions for the formation of intra-chain disulfide bonds and dimerization of subunits (Michaelis *et al.*, 1983; Hoffman & Wright, 1985; Manoil & Beckwith, 1985). The periplasmic enzyme is active regardless of whether it is transformed into the mature form or remains as translocated precursor with uncleaved signal peptide due to mutation. The precursor that is not translocated and therefore stays in the cytoplasm has no such activity (Boyd *et al.*, 1987). Enzymatic activities of the mature protein and the translocated precursor can be individually estimated directly in the gel after non-denaturing electrophoresis and gel staining, since mature protein isoforms and the translocated precursor have different electrophoretic mobilities. Mature phosphatase is known to be found under these conditions as isoforms I, II and III (Nesmeyanova *et al.*, 1981), resulting from proteolytic modification (Nakata *et al.*, 1987), and multimer forms (m; Figure 4), whereas the translocated precursor with uncleaved signal peptide is found at the top of the gel (p; Figure 4), in forms probably resulting from protein aggregation due to the presence of hydrophobic signal peptide (Karamyshev *et al.*, 1994b; Nesmeyanova *et al.*, 1994). It should be noted that non-translocated alkaline phosphatase precursor under the above conditions is not revealed due to the lack of enzymatic activity. The above properties of alkaline phosphatase allowed us in the present study to use the analysis of multiple molecular enzyme forms for elucidation of cell localization of mutated phosphatase precursors. Unprocessed mutant protein

THIS PAGE BLANK (USPTO)

with the substitution of Val for Ala(-1) was used as control (Figure 4); its enzymatic activity, translocation across the cytoplasmic membrane and periplasmic localization were shown previously using subcellular fractionation (Karamyshev *et al.*, 1994b; Nesmeyanova *et al.*, 1994). A completely non-translocated precursor due to impairment of the hydrophobic core (Michaelis *et al.*, 1986), e.g. protein with the substitution of Glu for Ala(-13), was used as another control. No active enzyme forms were found in this case (Figure 4). Analysis of mutant proteins obtained in this study showed that all proteins displaying a complete block of processing in pulse-chase experiments; for example those containing Pro(+1), Phe(-1), and Leu(-1) (Figures 2 and 3), were found by non-denaturing electrophoresis (Figure 4) at the top of the gel, in forms that are characteristic of the translocated precursors possessing enzyme activity. The electrophoretic pattern of these proteins is the same as that observed for the control unprocessed protein with Val(-1) (Figure 4). Mutant alkaline phosphatases with a slowing down of the processing rate, whose precursors were found even 60 minutes after pulse labeling (see Figures 2 and 3), contain both isoforms of the mature protein and forms of the translocated precursor (Cys(+1), Pro(-1), Cys(-1), Ala(-2), Pro(-2), Cys(-2), Phe(-5), Lys(-5), Tyr(-5), Figure 4). This also indicates that the unprocessed precursor is translocated across the cytoplasmic membrane. However, in some cases, e.g. in protein with Lys(-5), the relative content of active precursor is lower than could be expected, on the basis of the results of pulse labeling (Figure 3). It is quite probable that not only processing but also translocation of this protein is disturbed. Interestingly, protein with Phe(-5) showed a much greater amount of active precursor than could be expected from the results of pulse

labeling. Probably, such unprocessed precursor can be accumulated in the cells during cultivation. In the case of mutant proteins whose precursors are slowly but completely turned into the mature form (Figure 3), only mature phosphatase isoforms were found, as expected. As an example, Figure 4 shows multiple molecular enzyme forms of the mutant protein with Ala(-5).

On the basis of data obtained, the impairment of processing of most mutant proteins could be conditioned by the effect of amino acid substitutions on signal peptide cleavage by leader peptidase due to changes in the structure of protein as a substrate for this enzyme.

Discussion

E. coli alkaline phosphatase is a typical secreted protein with a canonical signal peptide (Inouye *et al.*, 1982). Amber-suppressor mutagenesis used in this study allowed us to obtain a wide range of mutant alkaline phosphatases with amino acid substitutions in the signal peptide cleavage site and to screen their effect on the processing. In order to understand the impact of these mutations, we also analyzed the distribution of residues in the cleavage region of 151 proteins from Gram-negative bacteria (Table 1). The results obtained give new experimental confirmation of the -3, -1 rule revealed earlier by statistical evaluation of the known primary structures of signal peptides (von Heijne, 1983a,b, 1986a; Perlman & Halvorson, 1983). It is quite obvious that the interaction of signal peptidase and the cleavage site of the alkaline phosphatase signal peptide proceeds most efficiently in the presence of Ala at position -1. It is this amino acid that is found at the signal peptide C terminus in most secreted proteins of Gram-

Table 1. Amino acid count for 151 known cleavage sites of proteins from Gram-negative bacteria

Amino acid	-6	-5	-4	-3	-2	-1	+1
Ala	46	25	27	99	12	139	64
Arg	0	1	1	1	0	0	2
Asn	4	5	7	0	9	0	4
Asp	0	0	0	1	1	0	14
Cys	5	1	0	1	0	0	3
Gln	1	9	6	2	15	0	11
Glu	0	1	1	0	0	0	18
Gly	5	17	15	0	0	8	7
His	0	4	2	0	17	0	1
Ile	3	2	1	2	2	0	0
Leu	5	13	7	5	24	0	3
Lys	0	0	1	0	1	0	8
Met	5	2	3	1	11	0	0
Phe	5	8	5	1	21	0	0
Pro	7	10	20	1	0	0	0
Ser	50	24	35	14	9	4	5
Thr	13	23	11	5	8	0	6
Trp	0	1	0	0	5	0	0
Tyr	0	1	2	1	11	0	2
Val	2	4	7	17	5	0	3

* The collection of proteins was extracted as described in Materials and Methods.

THIS PAGE BLANK (USPTO)

negative bacteria (Table 1), which is in agreement with a complete block of the processing resulting from most substitutions of this amino acid. The data obtained also give the first evidence of a higher conservation of amino acid residues at position -1 of signal peptides of Gram-negative bacteria as compared with that of eukaryotic signal peptides (Folz *et al.*, 1988). Gln and Cys found at this position of eukaryotic signal peptides (Nielsen *et al.*, 1997) appreciably (or completely) inhibited processing when in the same position in alkaline phosphatase. The specificity of signal peptidase of Gram-negative bacteria at position -1 seems to be as narrow as that of the thylakoid processing peptidase (Dalbey & von Heijne, 1992). The experimentally revealed lower regularity of position -3 is in accord with the fact that Val, Ser and other amino acid residues are found at this position as well as the most frequent, Ala. Moreover, the residues banned by the -3, -1 rule, such as Phe, Tyr, Arg, Asp and Pro, are occasionally found at this position in Gram-negative bacteria.

Position +1, which is also a constituent of the cleavage site, is not so conserved. It is known that all 20 amino acids are found at this position in the precursors of eukaryotic secreted proteins (Watson, 1984); however, Trp, Ile, Phe, Met and Pro were not found at position +1 of secreted proteins of Gram-negative bacteria and Cys was found at this position only in lipoproteins (Watson, 1984; von Heijne, 1986a). Our results show that only the introduction of Pro instead of Arg(+1) prevents signal peptide cleavage, which agrees with the data on the processing of the M13 procoat protein (Shen *et al.*, 1991) and the maltose-binding protein (Barkocy-Gallagher & Bassford, 1992). The latter work showed that the mechanism of inhibition of the processing of the maltose-binding protein with Pro(+1) differs from that initiated by amino acid substitutions at positions -3 and -1 of the signal peptide. If in the latter case the mutations disturb recognition of the cleavage site by signal peptidase, the presence of Pro(+1) transforms such a protein into an uncleavable substrate for leader peptidase but does not disturb the recognition of the mutant protein by signal peptidase. The above mutant protein acts as a competitive inhibitor of leader peptidase, which results in the accumulation of a number of secreted preproteins in the cells. It is quite probable that the mechanism of inhibition of the processing of the mutant alkaline phosphatase with Pro(+1) is similar to the one described above.

We also revealed that the introduction of Cys into position +1 results in impairment of the processing of a certain pool of precursor molecules: while most molecules of the precursor are rapidly processed already after six seconds, the rest of them remain unprocessed even after one hour. A similar effect was observed in proteins with Cys(-2) and Cys(-1). These data point to the fact that a certain number of precursor molecules of the above mutant proteins are impervious to attack by signal peptidase. It is quite probable that some pro-

tein modification, e.g. the formation of disulfide bonds with nearby protein molecules or modification of the residue Cys as in lipoproteins, may result in incomplete processing. Such a modification probably does not affect Cys residues at position -3, -4 or -5, and the appropriate mutant proteins are efficiently processed. This allows us to suggest that amino acid residues in the signal peptide cleavage site up to -2 are exposed to the periplasm, while the other part of the signal peptide is screened by the membrane.

Furthermore, the present study shows that recognition by signal peptidase depends on the nature of the amino acid residues not only at positions -3 and -1, but also to a lesser extent at positions -2 and -5, with their size being of key importance. Thus, a decrease in processing was found only in the presence of small amino acid residues (Gly, Ala and Cys) at position -2. This is in line with the previous statistical evaluation by von Heijne (1986a) and our analysis (Table 1). At the same time it is worth mentioning that although Ala occurs fairly frequently at position -2, its presence at this position in alkaline phosphatase partially decreases the processing rate. The explanation of the decline in processing rate may be in the potential mis-recognition of Gly and Ala at position -2 by the peptidase pockets predestined for the -1 and -3 residues. The effect of Cys may have another reason, which is described above.

Among 12 residues, substituting for Pro at position -5, Lys, Tyr, Leu, Ala and Phe decreased the efficiency of the processing. The effect of Gly and His was less significant, and Gln, Glu, Cys and Ser had no effect. Interestingly, our statistical analysis shows that despite the observed inhibiting effect of Phe, Leu, Tyr, Ala and Gly at position -5 of the alkaline phosphatase precursor (Figure 3), these residues are found at this position in other known signal sequences. A more detailed analysis of the known cleavage sites revealed that the presence of the bulky residues Phe, Leu and Tyr might be connected with the context of the cleavage site. For example, we found that the bulky residues (Trp, Tyr, Phe, Met and Leu) at position -5 occur more frequently when the cleavage site has Gly(-1) or Val(-3) (11% of cleavage sites have Val at the -3 position, but with a bulky residue at position -5, Val is found in 28% of sequences; in the case of Gly(-1) these values are 5% and 12%, respectively). Probably, a specific context of amino acid residues in the signal peptide cleavage region is needed for efficient processing, which may explain a discrepancy between the observed inhibition of the PhoA processing and the occurrence of amino acid residues in the known sequences of this region. Additional experimental investigations should be designed to examine this suggestion.

Interestingly, our statistical analysis also shows that the bulky residues Trp, Phe, Tyr, Met, Leu and Ile are almost absent at position -5 of the c-regions of Gram-positive bacteria. This suggests that the observed effect of bulky residues at pos-

THIS PAGE BLANK (USPTO)

ition -5 may also be found in Gram-positive bacteria.

It is known that the net charge distribution at the N and C termini of the signal peptide and a few amino acid residues of the mature protein's N terminus is important for secretion (von Heijne, 1986b). Since the introduction of Lys(-5) may change the above charge distribution, we cannot exclude an effect of this mutation on protein translocation across the membrane and therefore on the processing rate. However, the effects of Phe, Leu, Tyr, Ala and Gly do not fit in with this explanation. These residues are hydrophobic, hence the reason for their effect on processing may lie in a non-specific increase in the cleavage region's hydrophobicity. Laforet & Kendall (1991) suggested that the level of hydrophobicity of amino acid residues in the cleavage region of the signal peptide might be important for processing. However, Shen *et al.* (1991) and Barkocy-Gallagher *et al.* (1994) showed that the hydrophobicity of the cleavage site did not affect the processing. Our analysis of the known signal sequences also shows that although the cleavage regions are usually less hydrophobic than the central h-regions, the variability of their hydrophobicity is high. Some signal peptides, e.g. D-galactose-binding periplasmic protein (Mahoney *et al.*, 1981) and α -amylase (Schneider *et al.*, 1992) have a more hydrophobic c-region than the corresponding region of the mutant alkaline phosphatase signal peptide. Thus, a non-specific increase of the cleavage site hydrophobicity is probably not the reason for the inhibiting effect of the above amino acid substitutions.

Amino acid residues at position -5 having an inhibiting effect on alkaline phosphatase processing were revealed to contain either bulky or small side-chains. This may be important for the complementary binding of residues at position -5 with signal peptidase. Thus, all this suggests that the effect of Phe, Leu, Tyr (bulky) and Ala, Gly (small) residues at position -5 may be due to their size rather than their hydrophobicity.

Assuming that the cleavage site of secreted proteins must have one unique conformation competent for interaction with the active center of leader peptidase, we performed a stereochemical analysis of the region ranging from the -5 to +1 residues in order to establish this conformation. The absence of glycine preference at any position within the -5 to +1 region indicates that the residues of this fragment most probably have α and (or) β -conformations rather than glycine-specific α_L or β^* conformations (symbols α , α_L , β , and β^* denote, respectively, residue backbone conformations close to the right and left α -helical conformations as well as β -conformation and its mirror-symmetrical form). The known -3, -1 rule (von Heijne, 1983) predicts the interfacial (i.e. directed towards the active center of signal peptidase) orientations of side-chains of the residues at positions -3 and -1. At the same time, most amino acid substitutions at positions -2 and +1 have almost no effect on the

processing, which could be explained by an exterior orientation of these residues in the signal peptide-peptidase complex. This suggests an extended β -conformation of the region that ranges from the -3 to -1 residues (Dalbey & von Heijne, 1992). The conformation of the preceding region ranging from residues -6 to -4, was most frequently considered to be a β -turn (Barkocy-Gallagher *et al.*, 1994). Previously it was suggested that the presence of Pro, which is capable of breaking the α -helix conformation at a certain distance from the processing site, is required to form the β -turn to provide a curve in the polypeptide chain at the region between the signal peptide hydrophobic domain and the mature polypeptide N terminus (Perlman & Halvorson, 1983). However, we show here that when Pro residues are simultaneously present at both the -5 and -4 positions, alkaline phosphatase is processed properly. Only the β -conformation is sterically allowed for the first Pro of this Pro-Pro tandem (Williamson, 1992). Therefore, this suggests a β -conformation for the residue at position -5 and, as a consequence, casts doubt on the β -turn at the -6 to -4 region.

Two conformations, $\beta\alpha\beta\beta$ and $\beta\beta\beta\beta$, of the -5 to -1 region fulfil the condition that the -3, -2, -1 and -5 residues must have a β -conformation (Figure 5). An extended region of the -3, -2 and -1 residues is superimposed in Figure 5 in order to demonstrate the difference between the -5, -4 side-chain orientations in these two conformations. Our mutation analysis showed that mutations at position -4 have no effect on the processing. This suggests an exterior orientation of the -4 side-chain in the signal peptide-peptidase complex. Most residues at the -5 position that inhibit the processing have either bulky (Phe, Tyr, Leu) or

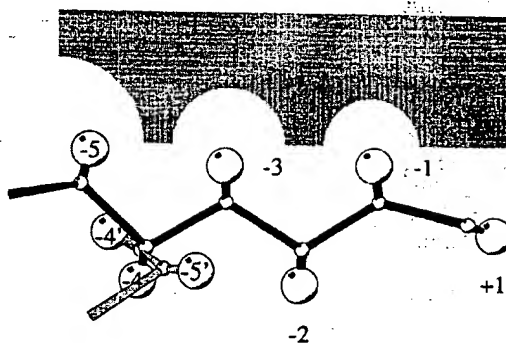


Figure 5. Ball-and-stick representations of two possible conformations of the processing region ranging from the -5 to +1 residues of a signal peptide. Shadow contours over the peptide fragments schematically designate the potential active site of the signal peptidase. C_α atoms of the polypeptide backbone are shown as small balls, C_β atoms of the side-chain groups are shown as large balls. The regions of the -3, -2, -1 residues having identical conformation are superimposed. The sticks showing peptides with $\beta\beta\beta\beta$ conformation and with $\beta\alpha\beta\beta$ conformation are black and gray, respectively.

THIS PAGE BLANK (USPTO)

small (Ala, Gly) side-chains. This may imply that the side-chain at position -5 is directed towards the binding center of the "receptor" molecule. If peptidase has a cavity for medium-sized residues of the signal peptide at position -5, it should cause energetically unfavorable steric tensions in the substrate-enzyme complex in the case of Phe, Tyr and Leu, or a lack of close packing in the case of Ala and Gly. The β -turn model (Barkocy-Gallagher *et al.*, 1994) with the conformation $\alpha\beta\beta\beta$ predicts the -4 residue to be closer to the interface with peptidase than the -5 residue. Thus, in contrast to the "extended" conformation, the β -turn model cannot provide any simple explanation of the effects of our mutations at positions -4 and -5.

As the statistical analysis shows, the occurrence of bulky residues at position -5 is often accompanied by the presence of Gly at position -1 or Val at position -3. In our structural models this relationship may be explained by the assumption that the small Gly(-1) (or bulky Val(-3)) allows the bulky -5 residue to come out from the receptor cleft and to relieve the tension that is usually present (Figure 5). We do not have enough evidence to choose unconditionally one of the two conformations shown in Figure 5. However, the location of residues at position -5 in the $\beta\alpha\beta\beta$ conformation is further from the putative enzyme surface compared with those in the $\beta\beta\beta\beta$ conformation. Thus, among several theoretically possible conformations this structure is more consistent with all the experimental and statistical data obtained in this study.

The hydrophobic h-region of the signal peptide is probably buried in the non-polar membrane with the $N_{in}-C_{out}$ orientation and has an α -helical conformation (Pugsley, 1993; Kajava *et al.*, 1991). The hydrophobicity of the c-region and evidence that the active site of the signal peptidase probably faces the periplasm (Dalbey, 1991) suggest a periplasmic location for the cleavage fragment. To visualize possible spatial arrangements of the c-region relative to the h-region and membrane, we modeled several structures, which have α -helical conformation from the N terminus to the -7th residue, followed by the c-region. A six-residue size was chosen for the c-region as the most frequently occurring in signal peptide sequences. Four possible arrangements were generated by a combination of the α and β conformations at positions -6 and -4 (Figure 6), while residues -5, -3, -2, -1 were fixed in β conformations. All arrangements fit our data on Cys mutants showing that residues at positions -2, -1, +1 are probably more exposed to the periplasm compared with residues -3, -4 and -5. The frequent occurrence of the helix-breaking Pro at positions -6 and -5 favors $\beta\beta\beta\beta$ and $\beta\alpha\beta\beta$ conformations (Figure 6a and b) as compared with $\alpha\beta\alpha\beta\beta$ and $\alpha\beta\beta\beta\beta$ conformations (Figure 6c and d). There are no data about the effect of h-region sequence on the processing that predict a distant arrangement of the signal

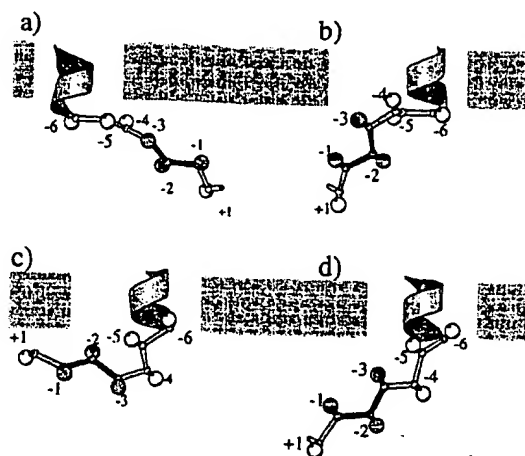


Figure 6. Schematic pictures of several possible arrangements of the α -helical h-region (ribbon representation), c-region (ball-and-stick) and the non-polar environment of the membrane (in gray). -3, -2, -1 region is outlined by black sticks and gray balls. Location of the leader peptidase active site can be imagined according to the assumption that -3 and -1 side-chains are directed into the binding pocket of the peptidase. a, $\beta\beta\beta\beta$; b, $\beta\alpha\beta\beta$; c, $\alpha\beta\alpha\beta\beta$; and d, $\alpha\beta\beta\beta\beta$, conformation of the c-region.

peptide α -helix from the c-region-binding site. If we compare $\beta\alpha\beta\beta$ and $\beta\beta\beta\beta$ conformations, the latter arrangement, being more distant, appears to be more probable. Thus this consideration also suggests an extended β -conformation for the c-region.

Materials and methods

Escherichia coli strains and plasmids

E. coli strains used in this study are listed in Table 2. Su^- strain E15 as well as strains MK376, MK377, MK378 and MK379 carrying *supD*, *supE*, *supF* and *supP* alleles, respectively, were used for amber suppressor mutagenesis (see Results). Strain Z85 was used as a host for vector and recombinant plasmids except for the plasmids with cloned amber suppressor tRNA genes (see below), where the XAC-1 strain was used. Vector plasmid p15SK(-) contains multiple cloning sites identical to pBluescript SK (Stratagene, USA), has the p15A *ori* of replication and the marker of resistance to chloramphenicol (Fischer & Hengstenberg, unpublished). Plasmid pPHO1 was the source of the *E. coli* alkaline phosphatase gene (Zozulya *et al.*, 1990). Plasmids pPHOA-wt, pPHOA-am(-5), pPHOA-am(-4), pPHOA-am(-3), pPHOA-am(-2), pPHOA-am(-1) and pPHOA-am(+1) were constructed here on the basis of vector p15SK(-). They carry the wild-type and mutant alkaline phosphatase genes cloned into the *HindIII*/*Bam*HI sites and having amber mutations at positions corresponding to the -5, -4, -3, -2, -1, +1 amino acid residues of this protein. Vector plasmid pGFIB1 with ColEI *ori* of replication and the marker of resistance to ampicillin and plasmids with the *E. coli* amber suppressor tRNA genes (Ala2,

THIS PAGE BLANK (USPTO)

Table 2. List of the *Escherichia coli* strains used

Strain	Genotype	Reference
Z85	<i>thi</i> Δ (<i>lac-proAB</i>) Δ (<i>srl-recA</i>) <i>hsdR</i> Tn10 F[<i>traD36 proAB laq1</i> ⁺ Δ M15]	Zaitsev <i>et al.</i> (1986)
E15	Hfr Δ <i>phoA8 fadL701 tonA22 garB10 ompF627 relA1 pit-10 spoT1 T2</i>	Bachmann (1987)
XAC-1	<i>ara</i> Δ (<i>lac-pro</i>) <i>nalA rif argE_{am} rpoB thi</i> F[<i>lacI373 lacZ_{U118am}</i>]	Normanly <i>et al.</i> (1986)
MK376	<i>ara</i> Δ (<i>lac-pro</i>) <i>supD gyrA metB argE_{am} rpoB thi</i> Δ <i>phoA8</i> Tn10	Karamyshev <i>et al.</i> (1994a)
MK377	<i>ara</i> Δ (<i>lac-pro</i>) <i>gyrA argE_{am} rpoB thi</i> Δ <i>phoA8</i> Tn10 (<i>cl⁺ glnV44 glnV89</i>)	Karamyshev <i>et al.</i> (1994a)
MK378	<i>ara</i> Δ (<i>lac-pro</i>) <i>supF gyrA metB argE_{am} rpoB thi</i> Δ <i>phoA8</i> Tn10	Karamyshev <i>et al.</i> (1994a)
MK379	<i>ara</i> Δ (<i>lac-pro</i>) <i>gyrA argE_{am} rpoB thi supP</i> Δ <i>phoA8</i> Tn10	Karamyshev <i>et al.</i> (1994a)

GluA, Phe, Gly1, HisA, ProH, Cys) cloned in the above vector (Kleina *et al.*, 1990) were provided by Dr G. Miller. Plasmids with cloned genes of amber-suppressor tRNAs of the phage T5 specific to Lys and Gly were constructed by Ksenzenko *et al.* (unpublished).

DNA techniques

Recloning of the *E. coli* alkaline phosphatase gene, analysis of recombinants obtained, isolation of a single-stranded phage DNA, plasmid DNA, phosphorylation of oligonucleotides and DNA sequencing were performed according to the protocols of Sambrook *et al.* (1989). Oligonucleotide-directed mutagenesis was carried out as described (Karamyshev *et al.*, 1994a). The mutagenic primers used in the work were as follows: 5'-CGT-TACTGTTTACCTAGGTGACAAAAGCC-3'; 5'-CTGT-TACCCCTTAGACAAAAGCCC-3'; 5'-TTACCCCTG-TGTAGAAAAGCCCGAC-3'; 5'-CCCCTGTGACATAG-CCCGGACACC-3'; 5'-CCCTGTGACAAAATAGCG-GACACCAG-3'; 5'-GTGACAAAAGCCTAGACACCA-GAAATGC-3'. They were used for the introduction of amber stop codons into positions corresponding to the -5, -4, -3, -2, -1 and +1 amino acid residues of the *E. coli* alkaline phosphatase, respectively. Mutations were confirmed by DNA sequencing.

Alkaline phosphatase maturation

Pulse-chase experiments to analyze alkaline phosphatase maturation were performed according to the method of Michaelis *et al.* (1986) with some modifications. *E. coli* cells grown to the mid-log phase in mineral medium (Torriani, 1966) with 1 mM K₂HPO₄ were harvested, washed, and incubated for 30 minutes in the same medium without orthophosphate for alkaline phosphatase induction. The cells were labeled with L-[³⁵S]methionine (50 μ Ci/ml) for 60 seconds and chased for 0.1, 1.0, 5.0 or 60.0 minutes by addition of unlabeled methionine to a final concentration of 0.5%. Alkaline phosphatase and its precursor were immunoprecipitated using affinity-purified rabbit antibodies against alkaline phosphatase and separated by 10% SDS-PAGE according to the method of Laemmli (1970), followed by autoradiography. Quantitation of proteins was performed using LKB UltroScan laser densitometer and the relative quantity of mature alkaline phosphatase from the total amount of the mature protein and its precursor was calculated with adjustments of additional methionine residues in the precursor.

Alkaline ph sptatase isof rms

A separate colony of *E. coli* was inoculated in 2 ml of mineral medium (Torriani, 1966) supplemented with

0.1 mM K₂HPO₄ and the corresponding antibiotics (ampicillin, 100 μ g/ml; chloramphenicol, 25 μ g/ml). Cells were cultured at 37°C under intensive aeration for 16 hours, harvested by centrifugation, and incubated in lyzing buffer (20 mM Tris-HCl (pH 7.5), 10 mM EDTA, 50 mM sucrose, 1 mg/ml lysozyme) for 15 minutes at 0°C. Samples separated by centrifugation from the cell debris were analyzed by non-denaturing electrophoresis in 7.5% PAGE (Davis, 1964). Staining of the alkaline phosphatase isoforms by their enzymatic activities was performed by treatment of the gel with α -naphthyl phosphate (Sigma, N-7255) and Fast Red Dye TR (Chemapol, Czechoslovakia) (Lojda *et al.*, 1979).

Statistical and stereochemical analysis

The collection of 151 exported proteins from Gram-negative bacteria with known cleavage sites was extracted from SwissProt 32.0 (Bairoch & Apweiler, 1996) using the Sequence Retrieval System software (<http://www.ebi.ac.uk/srs/srsc/>). It does not include protein sequences with more than 80% identity. Anomalous signal sequences (whose lengths of the hydrophobic core did not fall to the range between 7 and 17 residues), lipoproteins and proteins of pili were also excluded. The collection of exported proteins from Gram-positive bacteria was taken from sequence database SIGNALP (<ftp://virus.cbs.dtu.dk/pub/signalp/>) (Nielsen *et al.*, 1997).

To construct the structural models, conformations of the six-residue fragment of the c-region were adjusted manually by varying torsion angles of the polypeptide chain using TURBO-FRODO (Roussel & Combilian, 1989). The conformations were adjusted with regard to *a priori* requirements for the orientation of side-chains. The images in Figures 5 and 6 were generated by MOLSCRIPT (Kraulis, 1991).

Acknowledgments

We are grateful to J. H. Miller for kindly providing us with the collection of *E. coli* amber suppressors, R. Fischer for providing the plasmid vector p15SK(-) and M. G. Shlyapnikov for oligonucleotide synthesis. This study was supported by the Swiss National Science Foundation (grant 7SUPJ048349), the Russian Foundation for Basic Research (grant 96-04-48048), the Russian State Scientific Program "Modern methods of biological engineering", and was also made possible in part by an Award of the US Civilian Research and Development Foundation for the Independent States of the FSU (CRDF Award: RBI-288).

THIS PAGE BLANK (USPTO)

References

- Bachmann, B. I. (1987). Derivations and genotypes of some mutant derivatives of *Escherichia coli* K12. In *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology* (Neidhardt, F. C., ed.), pp. 1190–1219, American Society for Microbiology, Washington, DC.
- Bairoch, A. & Apweiler, R. (1996). NAR, the SwissProt protein sequences data bank and its new supplement. *TREMBL*, 24, 21–25.
- Barkocy-Gallagher, G. A. & Bassford, P. J., Jr. (1992). Synthesis of precursor maltose-binding protein with proline in the +1 position of the cleavage site interferes with the activity of *Escherichia coli* signal peptidase I *in vivo*. *J. Biol. Chem.* 267, 1231–1238.
- Barkocy-Gallagher, G. A., Cannon, J. G. & Bassford, P. J. (1994). β -turn formation in the processing region is important for efficient maturation of *Escherichia coli* maltose-binding protein by signal peptidase I *in vivo*. *J. Biol. Chem.* 269, 13609–13613.
- Boyd, D., Guan, C.-D., Willard, S., Wright, W., Strauch, K. & Beckwith, J. (1987). Enzymatic activity of alkaline phosphatase precursor depends on its cellular location. In *Phosphate Metabolism and Cellular Regulation in Microorganisms* (Torriani-Gorini, A., Rothman, F. G., Silver, S., Wright, A. & Yagil, E., eds), pp. 89–93, American Society for Microbiology, Washington DC.
- Chou, P. Y. & Fasman, G. D. (1978). Empirical predictions of protein conformation. *Annu. Rev. Biochem.* 47, 251–276.
- Dalbey, R. E. (1991). Leader peptidase. *Mol. Microbiol.* 5, 2855–2860.
- Dalbey, R. E. & von Heijne, G. (1992). Signal peptidase in prokaryotes and eukaryotes: a new protease family. *Trends Biochem. Sci.* 17, 474–478.
- Davis, B. J. (1964). Disc electrophoresis-II: method and application to human serum proteins. *Ann. NY Acad. Sci.* 121, 404–427.
- Fikes, J. D. & Bassford, P. J., Jr. (1987). Export of unprocessed maltose-binding protein to the periplasm of *Escherichia coli* cells. *J. Bacteriol.* 169, 2352–2359.
- Fikes, J. D., Barkocy-Gallagher, G. A., Klapper, D. G. & Bassford, P. J., Jr. (1990). Maturation of *Escherichia coli* maltose-binding protein by signal peptidase I *in vivo*. *J. Biol. Chem.* 265, 3417–3423.
- Folz, R. J., Nothwehr, S. F. & Gordon, J. I. (1988). Substrate specificity of eukaryotic signal peptidase. Site-saturation mutagenesis at position -1 regulates cleavage between multiple sites in human pre (Δ pro) apolipoprotein A-II. *J. Biol. Chem.* 263, 2070–2078.
- Hoffman, C. S. & Wright, A. (1985). Fusions of secreted proteins to alkaline phosphatase: an approach for studying protein secretion. *Proc. Natl Acad. Sci. USA*, 82, 5107–5111.
- Inouye, M., Barnes, W. & Beckwith, J. (1982). The signal sequence of alkaline phosphatase of *Escherichia coli*. *J. Bacteriol.* 149, 434–439.
- Izard, J. W. & Kendall, D. A. (1994). Signal peptides: exquisitely designed transport promoters. *Mol. Microbiol.* 13, 765–773.
- Kajava, A. V., Bogdanov, M. V. & Nesmeyanova, M. A. (1991). Stereochemical analysis of interaction of signal peptide with phospholipids at the initiation of protein translocation across the membrane. *J. Biomol. Struct. Dynam.* 9, 143–157.
- Karamyshev, A. L., Shlyapnikov, M. G., Khmel'nitsky, M. I., Nesmeyanova, M. A. & Ksenzenko, V. N. (1994a). Study of biogenesis and secretion of alkaline phosphatase and its mutant forms in *Escherichia coli*. I. Introduction of mutations into alkaline phosphatase gene. *Mol. Biol. (Moscow)*, 28, 150–157.
- Karamyshev, A. L., Kalinin, A. E., Tsfasman, I. M., Ksenzenko, V. N. & Nesmeyanova, M. A. (1994b). Study of biogenesis and secretion of alkaline phosphatase and its mutant forms in *Escherichia coli*. II. Effect of amino acid substitutions in the processing site and in N terminus of mature polypeptide chain on its biogenesis. *Mol. Biol. (Moscow)*, 28, 245–252.
- Kleina, L. G., Masson, J.-M., Normanly, J., Abelson, J. & Miller, J. H. (1990). Construction of *Escherichia coli* amber suppressor tRNA genes. II. Synthesis of additional tRNA genes and improvement of suppressor efficiency. *J. Mol. Biol.* 213, 705–717.
- Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* 24, 946–950.
- Kuhn, A. & Wickner, W. (1985). Conserved residues of the leader peptide are essential for cleavage by leader peptidase. *J. Biol. Chem.* 260, 15914–15918.
- Laemmli, U. K. (1970). Cleavage of structural protein during the assembly of the head of bacteriophage T4. *Nature*, 227, 680–685.
- Laforet, G. A. & Kendall, D. A. (1991). Functional limits of conformation, hydrophobicity, and steric constraints in prokaryotic signal peptide cleavage regions. Wild type transport by a simple polymeric signal sequence. *J. Biol. Chem.* 266, 1326–1334.
- Lajda, Z., Gossrau, R. & Schibler, T. H. (1979). *Enzyme Histochemistry. A Laboratory Manual*, Springer-Verlag, Berlin.
- Mahoney, W. C., Hogg, R. W. & Hermodson, M. A. (1981). The amino acid sequence of the D-galactose-binding protein from *Escherichia coli* B/r. *J. Biol. Chem.* 256, 4350–4356.
- Manoil, C. & Beckwith, J. (1985). *TrnphA*: a transposon probe for protein export signals. *Proc. Natl Acad. Sci. USA*, 82, 8129–8133.
- Michaelis, S., Inouye, H., Oliver, D. & Beckwith, J. (1983). Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli*. *J. Bacteriol.* 154, 366–374.
- Michaelis, S., Hunt, J. F. & Beckwith, J. (1986). Effects of signal sequence mutations on the kinetics of alkaline phosphatase export to the periplasm in *Escherichia coli*. *J. Bacteriol.* 167, 160–167.
- Nakata, A., Shinagawa, H. & Rothman, F. G. (1987). Molecular mechanism of isozyme formation of alkaline phosphatase in *Escherichia coli*. In *Phosphate Metabolism and Cellular Regulation in Microorganisms* (Torriani-Gorini, A., Rothman, F. G., Silver, S., Wright, A. & Yagil, E., eds), pp. 139–141, American Society for Microbiology, Washington, DC.
- Nesmeyanova, M. A., Motlokh, O. B., Kolot, M. N. & Kulaev, I. S. (1981). Multiple forms of alkaline phosphatase from *Escherichia coli* cells with repressed and derepressed biosynthesis of the enzyme. *J. Bacteriol.* 146, 453–457.
- Nesmeyanova, M., Karamyshev, A., Kalinin, A., Tsfasman, I., Badyakina, A., Khmel'nitsky, M., Shlyapnikov, M. & Ksenzenko, V. (1994). *Escherichia coli* alkaline phosphatase biogenesis: influence of overproduction and amino acid substitutions. In *Phosphate in Microorganisms: Cellular and Molecular Biology* (Torriani-Gorini, A., Yagil, E. & Silver, S.,

THIS PAGE BLANK (USPTO)

- eds), pp. 264-269, American Society for Microbiology, Washington, DC.
- Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.*, 1-6.
- Normanly, J., Masson, J.-M., Kleina, L. G., Abelson, J. & Miller, J. H. (1986). Construction of two *Escherichia coli* amber suppressor genes: tRNA^{Phe}_{CUA} and tRNA^{Cys}_{CUA}. *Proc. Natl Acad. Sci. USA*, 83, 6548-6552.
- Normanly, J., Kleina, L. G., Masson, J.-M., Abelson, J. & Miller, J. H. (1990). Construction of *Escherichia coli* amber suppressor tRNA genes. III. Determination of tRNA specificity. *J. Mol. Biol.* 213, 719-726.
- Perlman, D. & Halvorson, H. O. (1983). A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *J. Mol. Biol.* 167, 391-409.
- Pollitt, S., Inouye, S. & Inouye, M. (1986). Effect of amino acid substitutions at the signal peptide cleavage site of the *Escherichia coli* major outer membrane lipoprotein. *J. Biol. Chem.* 261, 1835-1837.
- Pugsley, A. P. (1993). The complete general secretory pathway in Gram-negative bacteria. *Microbiol. Rev.* 57, 50-108.
- Rosenblatt, M., Beadutte, N. V. & Fasman, G. D. (1980). Conformational studies of the synthetic precursor-specific region of preproparathyroid hormone. *Proc. Natl Acad. Sci. USA*, 77, 3983-3987.
- Roussel, A. & Cambillan, C. (1989). *Silicon Graphics Geometry Partner Directory* (Fall 1989), pp. 77-78, Silicon Graphics, Mountain View, CA.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory, New York.
- Schneider, E., Freundlieb, S., Tapio, S. & Boos, W. (1992). Molecular characterization of the MalT-dependent periplasmic α -amylase of *Escherichia coli* encoded by *malS*. *J. Biol. Chem.* 267, 5148-5154.
- Shen, L. M., Lee, J.-I., Cheng, S., Jutte, H., Kuhn, A. & Dalbey, R. E. (1991). Use of site-directed mutagenesis to define the limits of sequence variation tolerated for processing of the M13 procoat protein by the *Escherichia coli* leader peptidase. *Biochemistry*, 30, 11775-11781.
- Torriani, A. (1966). Alkaline phosphatase from *Escherichia coli*. In *Procedures in Nucleic Acid Research* (Cantoni, G. L. & Davis, R., eds), pp. 224-234, Harper and Row Publishers, New York.
- von Heijne, G. (1983). Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* 133, 17-21.
- von Heijne, G. (1986a). A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* 14, 4683-4690.
- von Heijne, G. (1986b). Net N-C charge imbalance may be important for signal sequence function in bacteria. *J. Mol. Biol.* 192, 287-290.
- Watson, M. E. E. (1984). Compilation of published signal sequences. *Nucl. Acids Res.* 12, 5145-5164.
- Williamson, M. P. (1992). The structure and function of proline-rich regions in proteins. *Biochem. J.* 297, 249-260.
- Zaitsev, E. N., Zaitseva, E. M., Bakhlanova, I. V., Gorelov, V. N., Kuzmin, N. P., Krykov, V. M. & Lantsov, V. A. (1986). Cloning and characterization of *recA* gene from *Pseudomonas aeruginosa*. *Genetics (Moscow)*, 22, 2721-2727.
- Zozulya, S. A., Obukhova, T. A., Shirokova, E. P. & Badalov, P. R. (1990). Plasmid expression vectors utilizing fragments of the *Escherichia coli* alkaline phosphatase gene. *Bioorg. Chem. (Moscow)*, 16, 1339-1347.

Edited by A. R. Fersht

(Received 20 August 1997; received in revised form 7 January 1998; accepted 9 January 1998)

THIS PAGE BLANK (USPTO)

Inefficient Membrane Targeting, Translocation, and Proteolytic Processing by Signal Peptidase of a Mutant Preproparathyroid Hormone Protein*

(Received for publication, September 27, 1994, and in revised form, November 7, 1994)

Andrew C. Karaplis†, Sung-Kil Lim§, Hisamitsu Baba¶, Andrew Arnold||, and
Henry M. Kronenberg

From the Endocrine Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114

A preproparathyroid hormone allele from a patient with familial isolated hypoparathyroidism was shown to have a single point mutation in the hydrophobic core of the signal sequence. This mutation, changing a cysteine to an arginine codon at the -8 position of the signal peptide, was associated with deleterious effects on the processing of preproparathyroid hormone to parathyroid hormone *in vitro*. To examine the biochemical consequence(s) of this mutation, proteins produced by cell-free translation of wild-type and mutant cRNAs were used in assays that reconstitute the early steps of the secretory pathway. We find that the mutation impairs interaction of the nascent protein with signal recognition particle and the translocation machinery. Moreover, cleavage of the mutant signal sequence by solubilized signal peptidase is ineffective. The consequence of this mutation on processing and secretion of parathyroid hormone is confirmed in intact cells by pulse-chase experiments following transient expression of the mutant protein in COS-7 cells. The inability of the mutant signal sequence, however, to interfere with the targeting and processing of other secreted proteins does not support obstruction of the translocation apparatus as the mechanism underlying the dominant mode of inheritance of hypoparathyroidism in this family.

Proteins destined for residence within membranes or for secretion contain hydrophobic amino-terminal sequences referred to as signal sequences (1). These sequences direct the nascent polypeptides bound to ribosomes to form a functional junction with rough endoplasmic reticulum (RER)¹ membranes, thereby assuring the translocation of the growing

polypeptide chain into the lumen of the endoplasmic reticulum and its subsequent cleavage by the luminal-localized signal peptidase enzyme (for review, see Ref. 2).

The nascent secretory protein is localized to the endoplasmic reticulum via a targeting apparatus consisting of the signal recognition particle (SRP) and its membrane-bound receptor on RER. The SRP binds to the signal sequence as it emerges from the large ribosomal subunit. This results in a transient delay or even arrest of translation (3) aimed in preventing premature folding of the precursor protein. When the SRP-ribosome complex encounters the SRP receptor or "docking protein," a series of reactions take place that result in the insertion of the nascent chain into the translocation site, release of SRP, resumption of translation, and initiation of translocation. GTP binding and its hydrolysis are required for these events to take place (4-6). As the nascent polypeptide transverses the translocation channel (7, 8) and emerges into the lumen of the endoplasmic reticulum, it is modified further by the signal peptidase enzyme complex that catalyzes the endoproteolytic cleavage of the signal sequence.

Only a small number of natural mutations in human signal sequences have been reported to have direct correlation with defective secretion and associated pathological states (9-11). We have described such a mutation in the signal peptide of one allele of preproparathyroid hormone (prepro-PTH) gene from a kindred with a form of familial isolated hypoparathyroidism (9). This is an inherited metabolic disorder characterized by hypocalcemia and hyperphosphatemia resulting from lack of biologically active circulating PTH, the major calcium-regulating peptide. In this family, the disorder was inherited as an autosomal dominant trait (12). The single point (T to C) mutation changed the codon at position -8 (signal peptide residues are numbered negatively starting from the site of cleavage toward the amino terminus) of the signal peptide of prepro-PTH from cysteine to arginine, thereby disrupting the hydrophobic core of the signal sequence (Fig. 1). Associated with this change was a dramatic impairment in the processing of the *in vitro* translated mutant prepro-PTH protein to pro-PTH by microsomal membranes (9).

Which step(s) of the early secretory process is affected by this mutation is not readily evident. Conceivably it could preferentially affect one or all of the steps involved, such as binding to SRPs, targeting to the RER, translocation through the membrane, and proteolytic processing by signal peptidase. In this report, we have systematically examined each of these steps using mutant and wild-type forms of *in vitro* translated prepro-PTH proteins by assaying their interaction with components of these various processes. Moreover, the consequence of this mutation on processing and secretion of PTH was examined in intact cells by transient expression of the mutant protein in COS-7 cells. Finally, we have used a co-transfection assay to

* This work was supported in part by National Institutes of Health Grant DK11794. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported by Medical Research Council of Canada Centennial Fellowship Award. To whom correspondence should be addressed: Div. of Endocrinology, Sir Mortimer B. Davis-Jewish General Hospital, McGill University, Lady Davis Inst. for Medical Research, Rm. 626, 3755 Côte Ste-Catherine Rd., Montréal, Québec, Canada H3T 1E2. Tel.: 514-340-8260 (ext. 4907); Fax: 514-340-7573.

§ Present address: Dept. of Internal Medicine, College of Medicine, Yonsei University, C. P. O. Box 8044, Seoul, Korea.

¶ Present address: Third Division of Internal Medicine, Kobe University School of Medicine, Kobe 650, Japan.

|| Present address: Laboratory of Endocrine Oncology, GRJ 1021, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114.

¹ The abbreviations used are: RER, rough endoplasmic reticulum; SRP, signal recognition particle; PTH, parathyroid hormone; PAGE, polyacrylamide gel electrophoresis.

THIS PAGE BLANK (USPTO)

[illegible]

FIG. 1. Signal peptide sequence of human prepro-PTH. Amino acids -25 to -1 constitute the signal peptide of wild-type and mutant (single and double) forms of human prepro-PTH. The 6 residues following the signal peptidase cleavage site (*arrowhead*) make up the pro sequence, and the 84 amino acids of mature PTH follow. Residues -16 to -5 (*underlined*) comprise the hydrophobic core of the signal peptide. The described patient's mutation (Cys → Arg at the -8 position; single mutant) and the additional substitution at the -10 position of the signal peptide (Ala → Arg; double mutant), are indicated by boldface. *Numbers above the amino acids* indicate their position relative to the signal peptidase cleavage site.

define the mechanism by which this specific mutation could cause clinical hypoparathyroidism in the presence of a second, apparently normal (9), PTH allele.

EXPERIMENTAL PROCEDURES

Materials—Restriction endonucleases, Klenow fragment of *Escherichia coli* DNA polymerase I, and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). Endoglycosidase H (endo- β -N-acetylglucosaminidase H) was from Boehringer Mannheim. Rabbit reticulocyte lysate, wheat germ extract, and canine pancreatic rough microsomes were purchased from Promega Corp. SRPs prepared from freshly excised canine pancreas and eluted from an aminopentyl-agarose column (13) were a generous gift from Reid Gilmore, University of Massachusetts, Worcester, MA.

Construction of Plasmids:—Plasmid pWT84 was prepared by inserting a *DdeI*-*XbaI* fragment of human prepro-PTH cDNA encoding prepro-PTH(1–84) within the *HindIII*-*XbaI* cloning sites of the mammalian expression vector pCDM8 (9). The single point mutation (T to C) was introduced at the –8 position of the signal peptide of prepro-PTH cDNA using oligonucleotide-directed mutagenesis (14), resulting in plasmid pSM84 (single mutant; see Fig. 1). Plasmid pSM84 was then further modified by two additional point mutations introduced into the –10 position of the signal peptide (GCA to CGA; Ala to Arg), resulting in plasmid pDM84 (double mutant; Fig. 1).

Plasmids pWT83, pSM83, and pDM83, all containing prepro-PTH(1-83) sequences without the termination codon, were constructed by subcloning a *HindIII-PstI* fragment containing the prepro sequences from plasmids pWT84, pSM84, and pDM84, respectively, into plasmid SP-PTH(*XbaI/Bal31/ClaI*/#6) comprising mature PTH(1-83) sequences without the termination codon (15).

In order to construct a plasmid encoding a protein of the same size as prepro-PTH but lacking a functional signal sequence, *Nco*I linkers were introduced into the *Sma*I site in the polylinker of SP-PTH(XbaI/Bal31/Clal#7) (15), and a 225-base pair *Hae*III-*Hind*III fragment derived from the same plasmid was ligated into the *Nco*I site. This introduced an ATG initiation codon followed by PTH sequences (51-82)-Ala in place of the signal sequence. This plasmid was then restricted with *Clal* and *Hind*III and ligated to a *Dpn*I-*Hind*III fragment isolated from plasmid SP-PTH (15), containing sequences encoding PTH(1-84). The resulting plasmid (pRM84, for random mutant signal sequence), encoded Met-PTH(51-82)-Ala-Ser-PTH(1-84).

To introduce an *N*-linked glycosylation site into the PTH-coding sequence of plasmids pWT84 and pSM84, synthetic oligonucleotides, encoding Ser-Asn-Gly-Ser-Gly-Glu-Gly-Val-Glu-Ser, were ligated into the unique *Taq*I site of the prepro-PTH coding sequence (the underlined sequence indicates the consensus sequence for *N*-glycosylation). The resulting plasmids, pWT84(G) and pSM84(G), differed from pWT84 and pSM84, respectively, only by the insertion of 9 amino acids between residues Ser-17 and Met-18 of the mature PTH protein.

In Vitro Transcription, Translation, and Analysis of Products—Plasmids were linearized and sense RNA strands were transcribed using

either T7 or SP6 RNA polymerase (15). Translation reactions in rabbit reticulocyte lysate and in wheat germ extract were performed according to the manufacturer's (Promega) procedure. Translation products were immunoprecipitated using affinity-purified goat anti-human PTH(1-34) antiserum and subjected to 15-20% continuous gradient SDS-polyacrylamide gel electrophoresis (PAGE). Autoradiography was performed after treating the gel with ENHANCE (DuPont NEN) as described previously (15). The identity of PTH-related peptides was established by amino-terminal radiosequence analysis (16).

Posttranslational Membrane Binding—Truncated cRNAs missing the termination codon were transcribed from plasmids pWT83, pSM83, and pDM83 and translated in the rabbit reticulocyte lysate cell-free system for 15 min at 24 °C. Aurintricarboxylic acid was added to 0.1 mM to inhibit translation initiation, and after another 15-min incubation, emetine was added (1 mM final concentration) to block peptide elongation. Incubation was continued for another 15 min, 4 eq (1 eq refers to 1 μ l of the original rough microsome preparation that has been adjusted to a concentration of 50 A_{280} units/ml; see Ref. 17) of microsomal membranes/25 μ l of translation mixture were added, and incubation continued for 15 min at 24 °C. Translation products were then incubated at 0 °C for 10 min. Insertion into the membranes was assessed by centrifugation of the membranes through either a physiological salt- or an EDTA-sucrose step cushion (15, 18). Supernatants containing membrane-free peptides and pellets were immunoprecipitated and subjected to SDS-PAGE analysis, as described (15).

Protease Protection Assay—To assay for co-translational translocation, wild-type and single-mutant cRNAs were translated in rabbit reticulocyte lysate cell-free system in the presence of canine pancreas microsomal membranes. After a 15-min incubation at 24 °C, proteinase K (20 μ g/ml final concentration) was added to the translation reaction mixtures either alone or in combination with Triton X-100 (1%), and incubation was continued for a further 60 min on ice. After inactivation of the protease with phenylmethylsulfonyl fluoride (final concentration, 2 mM), radiolabeled proteins were subjected to immunoprecipitation and SDS-PAGE analysis.

Endoglycosidase H Digestion—Transcribed products of pWT84(G) and pSM84(G) were translated in the absence or presence of microsomal membranes, immunoprecipitated, and treated with endoglycosidase H (10 milliunits) as described previously (19). Following overnight digestion, bovine serum albumin was added as a carrier (final 0.25%), and the samples were precipitated by ice-cold trichloroacetic acid (15% final concentration). Following centrifugation, the precipitates were resuspended in 0.1 M Tris-HCl, pH 7.4, and prepared for SDS-PAGE analysis as usual.

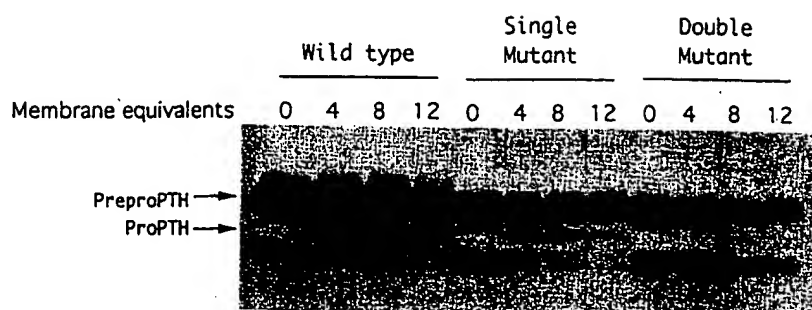
Signal Peptidase Assay—EDTA-stripped, nuclease-treated rough microsomes were prepared from canine pancreas as described (17). Aliquots of rough microsomes were resuspended by homogenization in ice-cold buffer (20 mM HEPES, pH 7.6, 50 mM NaCl) to a final concentration of 50 A₂₈₀ units/ml. Solubilization of signal peptidase with sodium deoxycholate was performed as described previously (20). Briefly, one volume of 10% (w/w) sodium deoxycholate was mixed with 19 volumes of membrane suspension. The resultant clear solution was centrifuged at 100,000 × g for 4 h, and the supernatant was frozen in liquid nitrogen. Wild-type and single-mutant prepro-PTH and prepro-PTH(G) cRNAs were translated in a wheat germ cell-free system, and translation products were used as substrates for detergent-solubilized signal peptidase (20). A typical 50-μl posttranslational cleavage assay contained 20 μl of the detergent extract, 20 μl of wheat germ translation mixture, and 10 μl of water. Posttranslational cleavage by signal peptidase was allowed to proceed at 25 °C for 90 min.

Pulse-Chase Analysis—COS-7 cells cultured in 12-well (22-mm diameter) plates in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum and antibiotics were transfected with plasmids pWT84, pSM84, and pDM84 using the DEAE-dextran protocol followed by 10% dimethyl sulfoxide shock (21). Five days later, cells were rinsed twice with Dulbecco's modified Eagle's medium without methionine and supplemented with 2% dialyzed fetal calf serum and then incubated for 15 min at 37 °C in the same medium containing [³⁵S]methionine (40 µCi/well). The cells were then washed once with 2 ml of Dulbecco's modified Eagle's medium (with methionine, supplemented with 10% fetal bovine serum), and then incubated further in the same medium for the times indicated. Conditioned media were saved for immunoprecipitation, and cells were lysed with 0.5 ml of lysis buffer (0.01 M Tris-HCl pH 7.2, 1% Triton X-100, 0.1% SDS, 1% sodium deoxycholate, 5 mM EDTA, 0.02% sodium azide, 0.15 M NaCl) with 50 µl of phenylmethylsulfonyl fluoride at a final concentration of 500 µg/ml. DNAs of the cell lysates were sheared with a 22-gauge needle, and the extracts were briefly centrifuged prior to immunoprecipitation.



THIS PAGE BLANK (USPTO)

FIG. 2. Processing of normal and mutant prepro-PTH. Autoradiogram of [3 H]leucine-labeled proteins derived from the translation of cRNAs transcribed from plasmids containing wild-type (pWT84), single mutant (pSM84), and double mutant (pDM84) prepro-PTH cDNA. Translation was performed in the rabbit reticulocyte lysate cell-free system in the absence (0 eq) or presence of increasing amounts (4, 8, and 12 eq) of canine pancreatic microsomal membranes.



Competition Studies—COS-7 cells were transfected with pWT84 or pSM84 either alone or in combination with pWT52 encoding for human prepro-PTH(1–52) (22). After the indicated times post-transfection, cells were labeled with [35 S]methionine for 15 min at 37 °C, as described above. Cells were then lysed and processed for immunoprecipitation and SDS-PAGE analysis.

RESULTS

Processing of Wild-type and Mutant prepro-PTH—To assess the processing of wild-type and mutant forms of prepro-PTH to pro-PTH, transcribed sense RNA strands were translated in a rabbit reticulocyte lysate cell-free system in the absence or presence of increasing concentrations of canine pancreatic microsomal membranes. In this co-translational assay, substitution of Cys → Arg at the –8 position of the prepro-PTH signal peptide (single mutant) resulted in impaired processing to pro-PTH as compared with the wild-type form (Fig. 2). The addition of a second charged residue within the hydrophobic core of the signal peptide (Ala → Arg at the –10 position, double mutant) completely abolished its proteolytic processing by rough microsomes. Maximal processing of wild-type and single-mutant precursors occurred when microsomal membranes were added to a final concentration of 4 eq/25 μ l of translation mixture.

Interaction with SRP—The impaired processing of the mutant prepro-PTH proteins by microsomal membranes may result from the inability of SRP to recognize efficiently the altered signal peptides. Binding of SRP to the signal sequence of nascent secretory proteins induces a site-specific elongation arrest of translation (3, 23). We, therefore, assessed the interaction of the mutant prepro-PTH signal sequences with SRP by examining the effect of exogenous SRP on inhibition of translation in a wheat germ cell-free system (Fig. 3). The mRNAs for pWT84, pSM84, pDM84, pRM84, and rabbit globin were translated in the presence (final concentration 0.02 and 0.04 A_{280} units/ml) or absence of exogenous SRP. The translation of the wild-type protein was inhibited significantly more than the translation of the mutant peptides by exogenous SRP. Moreover, translation of the single-mutant form was impaired to a greater extent than that of the double mutant. The addition of SRP did not affect the synthesis of globin (cytoplasmic protein with no signal peptide) and PTH-like protein with a random mutant signal peptide. These results suggested that the observed inhibition in processing of the mutant precursors is, at least, partly due to the inability of their signal sequences to interact effectively with SRP.

Translocation-competent Binding—To assess binding of the mutant prepro-PTH proteins to microsomal membrane components, truncated cRNAs for wild-type and mutant forms of prepro-PTH missing the termination codon were transcribed from plasmids pWT83, pSM83, and pDM83 and translated in the rabbit reticulocyte lysate cell-free system, thereby “freezing” the nascent proteins on ribosomes. After translation was complete, rough microsomes were added, and insertion into membranes was assessed following centrifugation of the reaction mixture through either a physiological salt- or an EDTA-

sucrose step gradient. Only nascent chains tightly bound to the translocation apparatus pellet with the membranes in the presence of high concentrations of EDTA (18). When membranes were added to the reaction mixture post-translationally, the propeptide (wild-type and single mutant) as well as the prepropeptide bands were seen in the precipitates (Fig. 4A). Moreover, peptides pelleting with the rough microsomes were resistant to extraction with EDTA, suggesting that translocation-competent binding of the ribosome-nascent chain complex to the membranes had taken place. This post-translational assay suggested that the mutant proteins can be targeted to microsomal membranes and bind in a translocation-competent manner. Although the mutant forms were able to engage the translocation apparatus, they did so less efficiently than the wild-type form. Equal loading of the various reaction mixtures was verified by examining corresponding supernatant fractions using SDS-PAGE analysis (Fig. 4B). Thus, the amount of nascent protein that sedimented with rough microsomes paralleled the affinity of SRP for the respective peptides (wild-type > single mutant > double mutant) likely reflecting their degree of interaction with SRP. As shown in Fig. 4A, processing of the mutant forms to pro-PTH was again markedly impaired when compared with the wild-type protein, even though translocation-competent binding had been achieved. These results suggested that additional processes in the early steps of the secretory pathway, i.e. translocation and/or interaction with signal peptidase, may be impaired by the mutant signal sequences.

Protease Sensitivity of Membrane-bound Nascent Chains—To determine the location of the nascent chains within the rough microsomes, we examined their sensitivity to digestion with proteinase K either in the absence or presence of Triton X-100. Protein products that are translocated to the lumen of the microsomal vesicles would be protected from digestion by proteolytic enzymes that are unable to enter these vesicles. As shown in Fig. 5, the processed peptides (pro-PTH) were protected by the membranes from proteolysis by proteinase K. The protection of these forms must have resulted from their sequestration into the lumen of the microsomal vesicles. This was confirmed by the addition of Triton X-100 to the reaction mixture, thereby, permeabilizing the membrane bilayer and allowing the protease to gain access to all protected polypeptides. The minor unprocessed single-mutant product that is protected from proteolysis may represent unprocessed nascent protein that has not fully translocated, yet is protected from proteolysis by the ribosomes and the tight ribosome-membrane junction required for translocation.

N-Glycosylation of a Modified PTH Sequence—Since the mutant signal sequence might be a poor substrate for signal peptidase, cleavage alone is an insufficient criterion for the localization of PTH peptides. The fact that a fraction of the mutant prepro-PTH was protected from proteolysis raises the question of how far into the translocation process the uncleaved precursor has progressed. Because glycosylation is restricted to the lumen of the endoplasmic reticulum, the addition of carbohy-

THIS PAGE BLANK (USPTO)

FIG. 3. Effect of SRP on translation. The cRNAs for wild-type (pWT84), single mutant (pSM84), double mutant (pDM84), random mutant (pRM84), and rabbit globin were translated in a wheat germ cell-free system in the absence (0) or presence of exogenous SRP (final concentration 0.02 (1) and 0.04 (2) A280 units/ml).

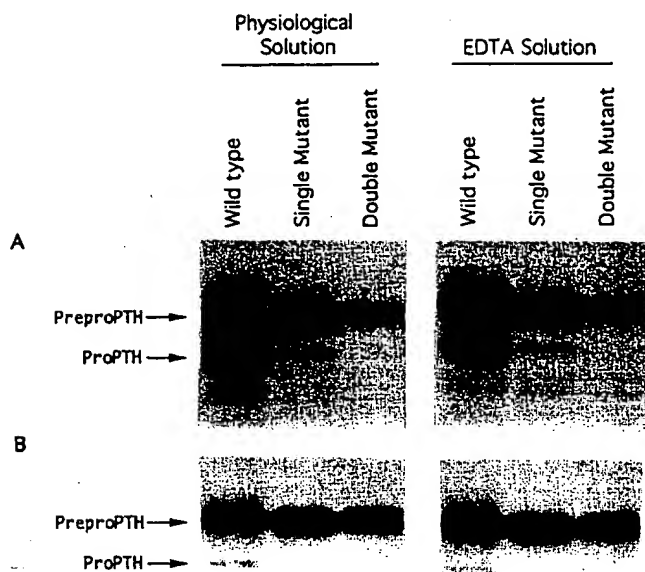
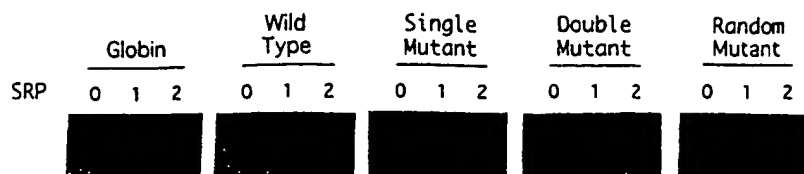


FIG. 4. Translocation-competent binding to microsomal membranes. Truncated cRNAs missing the termination codon were transcribed from plasmids encoding wild-type, single mutant, and double mutant cDNAs and translated in rabbit reticulocyte lysate system. Translocation-competent binding to microsomal membranes was assessed by centrifugation of the membranes through either a physiological salt- or an EDTA-sucrose step cushion. Radiolabeled proteins in the pellets (A) and corresponding supernatants (B) are displayed.

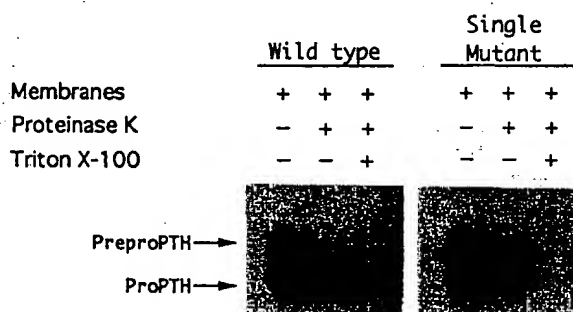


FIG. 5. Protease sensitivity of membrane-bound nascent chains. Wild type (pWT84) and single mutant (pSM84) cRNAs were translated in the rabbit reticulocyte lysate system in the presence (+) of dog pancreas microsomal membranes. After translation was completed, reactions were treated with proteinase K (20 μ g/ml) with (+) or without (-) the addition of 1% Triton X-100. Radiolabeled translation products were immunoprecipitated and analyzed by SDS-PAGE.

drate to the PTH sequence by microsomal membranes would constitute further evidence of translocation and would not directly require the presence of a suitable substrate for signal peptidase (19). Because the mature PTH protein does not have an N-linked glycosylation site, we engineered such a site 23 amino acid residues distal to the signal peptidase cleavage site.

Fig. 6 shows that upon translation of pWT84(G)-transcribed cRNA in reactions supplemented with microsomal membranes, two translational products with PTH immunoreactivity were seen that were not present in the absence of membranes. The smaller product migrated with an apparent molecular weight slightly greater than that of authentic pro-PTH and was there-

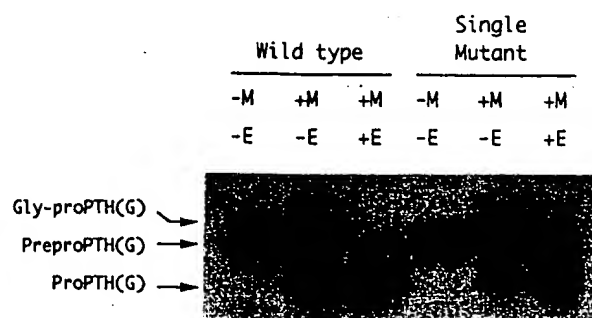


FIG. 6. Glycosylation of wild-type and mutant prepro-PTH(G). Plasmids pWT84(G) and pSM84(G) were transcribed *in vitro* and translated in the rabbit reticulocyte lysate cell-free system in the absence (-M) or presence (+M) of canine microsomal membranes. Translation products were immunoprecipitated and treated with (+E) or without (-E) endoglycosidase H.

fore felt to be pro-PTH(G). The second product, which migrated more slowly than prepro-PTH(G), was believed to be the glycosylated form of pro-PTH(G). This was confirmed by treatment with endoglycosidase H which, by removing carbohydrate on this peptide, shifted its position on an SDS-PAGE gel to that of pro-PTH(G). Interestingly, prepro-PTH(G) appeared to be processed much more efficiently by canine microsomal membranes than the unmodified form of the protein (compare Figs. 2 and 6), and this may simply reflect the influence of the extended length or the specific structure of the nascent chain.

Translation of pSM84(G)-transcribed cRNA in the presence of membranes resulted in the appearance of three PTH-immunoreactive products, two of which co-migrated with pro-PTH(G) and its glycosylated form, while the third was consistent with the unprocessed prepro-PTH(G) form. Again, the addition of 9 amino acids to the mature PTH molecule resulted in a more efficient cleavage of the signal sequence by microsomal membranes as compared with the unmodified form (see Figs. 2 and 6). Yet, the mutant signal sequence was once again processed less efficiently than the wild-type sequence, as indicated by the persistence of the unprocessed mutant prepro-PTH(G) form in the immunoprecipitated products. Once cleaved, however, pro-PTH(G) was glycosylated appropriately as confirmed by treatment of the reaction products with endoglycosidase H. The addition of carbohydrate to this moiety provides direct and independent evidence for its translocation, although significantly impaired, across the endoplasmic reticulum membranes. Furthermore, no larger glycosylated product was found that might have represented a protein that was translocated, glycosylated, and yet not cleaved by signal peptidase. Therefore, the uncleaved mutant prepro-PTH(G) was not delivered to the glycosylation machinery on the inner surface of the microsomal membranes.

Processing by Solubilized Signal Peptidase—To determine whether the single-mutant form of prepro-PTH is an unsuitable substrate for signal peptidase, we assessed proteolytic processing of the nascent protein in a translocation-independent assay. Following translation of wild-type and single-mutant prepro-PTH cRNAs in the wheat germ cell-free system, signal peptidase assays were performed by mixing aliquots of

THIS PAGE BLANK (USPTO)

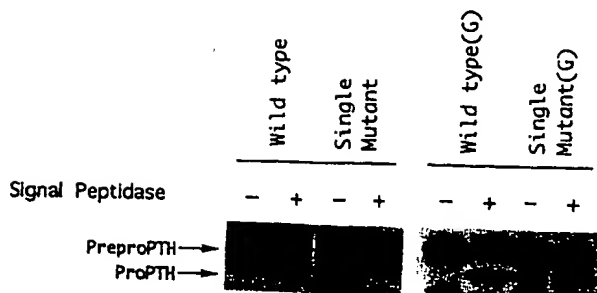


Fig. 7. Signal peptidase assay. Following translation of either pWT84- and pSM84- or pWT84(G)- and pSM84(G)-transcribed cRNAs in a wheat germ extract, signal peptidase assays were performed by mixing aliquots of translation mixture and signal peptidase prepared by directly solubilizing canine pancreatic rough microsomes.

translation mixture and signal peptidase prepared by directly solubilizing canine pancreatic rough microsomes. In this post-translational assay, no cleavage was detected for the mutant precursor, while a minor amount of processing was observed for the wild-type form of prepro-PTH (Fig. 7). Since the wild-type prepro-PTH is also an inefficient substrate for solubilized signal peptidase, an attempt was made to increase the sensitivity of the assay using as substrate the modified form of the prepro-PTH molecule containing an engineered glycosylation target site. RNAs encoding the wild-type and single-mutant form of prepro-PTH(G) were translated, and solubilized signal peptidase was added posttranslationally (Fig. 7). As found with intact membranes, wild-type and mutant forms of this modified PTH protein were processed more efficiently than their normal counterparts. Indeed, the mutant prepro-PTH(G) was processed by solubilized signal peptidase unlike its unmodified form, but the extent of cleavage was significantly less than that of the wild-type prepro-PTH(G). These results suggested that the mutation in the signal peptide makes the protein a less suitable substrate for signal peptidase.

Processing and Secretion of Prepro-PTH in COS-7 Cells—To confirm the cell-free data outlined above in intact cells, we examined the processing and secretion of mutant and wild-type forms of prepro-PTH from COS-7 cells transfected transiently with the corresponding expression plasmids. Five days following transfection, cells were pulse-labeled for 15 min with [³⁵S]methionine and then chased for the times indicated in Fig. 8 with medium containing cold methionine. The PTH species in cell extracts and conditioned media were immunoprecipitated and then resolved by SDS-PAGE. At the earliest time point examined (0 min), the predominant product immunoprecipitated in wild-type PTH-transfected cells was pro-PTH, while only barely detectable levels of prepro-PTH precursor were observed (Fig. 8A). Moreover, immunoreactive PTH was seen in the culture medium of these cells after 30 min of chase (Fig. 8B). The rate of processing seen with the wild-type sequence was in striking contrast to that observed with the mutant forms. In lysates from cells transfected with either mutant, the predominant band at the earlier time point corresponded with the prepro-PTH precursor with proteolytic cleavage to pro-PTH being dramatically reduced in efficiency. With the single mutant, only a small amount of pro-PTH was detected; no pro-PTH was found in cells expressing the double mutant. With the single-mutant form, a band corresponding to prepro-PTH was detectable even after 120 min of chase, consistent with the *in vitro* observed inefficient cleavage of this precursor to pro-PTH by microsomal membranes. Immunoreactive PTH was detectable in culture media of these cells after 30 min of chase but at substantially reduced levels as compared with the wild-type form. The double mutant form of prepro-PTH, although efficiently translated, was not processed to pro-PTH, nor was

immunoreactive PTH detectable in the culture media of cells transfected with this plasmid (data not shown). These results demonstrate in intact cells the inefficient processing of the mutant prepro-PTH molecules as compared with the wild-type form. The single-mutant prepro-PTH allows a small amount of normal processing, consistent with the results in the cell-free studies that showed that the single mutant molecule can engage the translocation apparatus, although less efficiently than normal. The double mutant prepro-PTH almost totally fails to engage the translocation apparatus in cell-free extracts and is not processed at all in intact cells.

Competition Experiments—A co-transfection assay in COS-7 cells was used to address the issue of dominant transmission of hypoparathyroidism in the family carrying the single-mutant prepro-PTH allele. The objective was to determine whether the inefficient processing of the mutant precursor can result in obstruction of the translocation apparatus and, thereby, to global defects in protein processing. For this study, an expression vector containing sequences encoding a truncated version of prepro-PTH, missing the last 32 residues (WT52), was used to provide a readily distinguishable variant of prepro-PTH with a normal signal sequence.

WT52 was co-transfected into COS-7 cells with either wild-type or mutant prepro-PTH expression vector. Over-expression of the mutant precursor for up to 10 days did not interfere with the processing of prepro-PTH(1-52) (Fig. 9). Although these results may simply reflect lack of sensitivity of the system, they do not support global interference with protein processing at the microsomal membrane level as a consequence of overexpression of the mutant prepro-PTH form.

DISCUSSION

Translocation from the cytoplasm into the endoplasmic reticulum, is a multistep process requiring a functional amino-terminal signal peptide. A signal sequence must perform effectively several distinct functions required for the efficient translocation of secreted proteins. These subfunctions include its recognition and binding to SRP, its interaction with membrane-bound components of the export machinery, opening the protein-conducting channels to initiate translocation, and appropriate presentation to the signal peptidase for cleavage.

Three domains have been identified as a common feature of eukaryotic signal sequences, and considered to be necessary for carrying out these functions: a positively charged NH₂ terminus, a central hydrophobic core of 10–15 amino acid residues, and a polar COOH-terminal region (24, 25). While the COOH-terminal region influences the efficiency and fidelity of signal peptidase cleavage, intactness of the hydrophobic region is indispensable for initiating translocation.

Two reported inherited mutations in the signal sequence of human secreted proteins, namely preprovasopressin (10), and preprofactor X (11), involve the COOH-terminal region of the respective signal peptides. Thus, a point mutation resulting in substitution of Arg for Gly at the -3 position of the factor X signal peptide (Factor X *Santo Domingo*) blocks cleavage by signal peptidase but does not interfere with targeting and translocation to the RER (11). Similarly, a naturally occurring substitution of Thr for Ala at the -1 position of the signal peptide of preprovasopressin results in central diabetes insipidus (10). This mutant protein, similar to the Factor X *Santo Domingo*, undergoes inefficient cleavage by signal peptidase, although targeting and translocation to the RER are not measurably affected.

The present study is the first to examine the effect of a naturally occurring substitution at the hydrophobic core of a signal peptide that results in human disease. Prepro-PTH, the precursor of PTH, contains a typical 25-residue amino-terminal

THIS PAGE BLANK (USPTO)

FIG. 8. Expression of wild-type and mutant forms of prepro-PTH in COS-7 cells. Plasmids pWT84, pSM84, and pDM84 were transiently transfected into COS-7 cells. Five days following transfection, the cells were pulse-labeled with [35 S]methionine for 15 min. At the indicated times following pulse-labeling, the media were removed, and cell extracts were prepared. Both cell extracts (A) and media (B) were immunoprecipitated using a PTH-specific antibody prior to SDS-PAGE analysis. *, sample not processed.

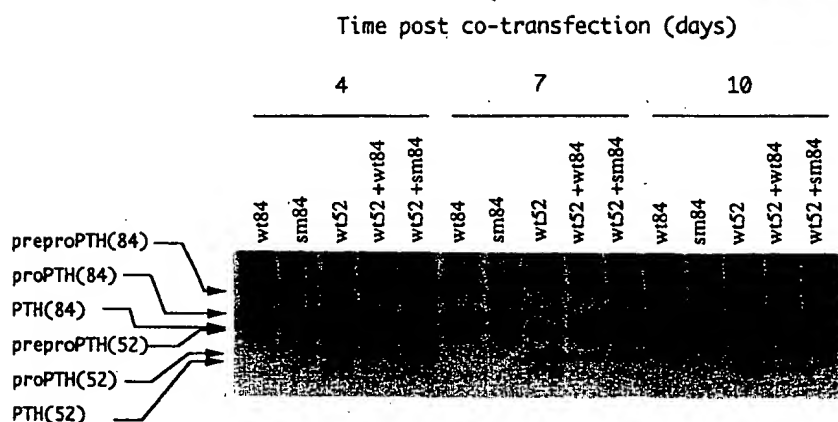
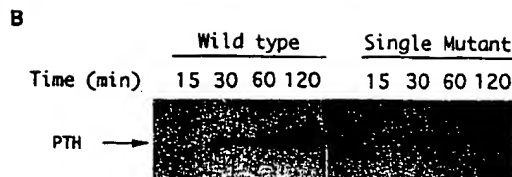
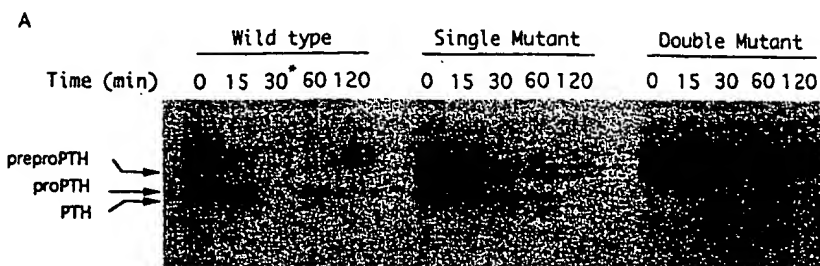


FIG. 9. Co-transfection of wild-type and mutant prepro-PTH. Plasmids pWT84 and pSM84 were transfected into COS-7 cells either alone or in combination with pWT52. After the indicated number of days posttransfection, cells were pulse-labeled with [35 S]methionine for 15 min. Cell extracts were then prepared and immunoprecipitated prior to SDS-PAGE analysis.

signal sequence followed by a 6-residue prospecific peptide and the mature hormone (residues 1–84; see Fig. 1). The hydrophobic core of the human prepro-PTH signal peptide is composed of 12 contiguous uncharged amino acids (residues –5 to –16 of the signal peptide).

In the present study, we have demonstrated that substitution of a charged amino acid, Arg for Cys, in the signal peptide hydrophobic core of prepro-PTH impairs co-translational translocation as well as posttranslational cleavage by isolated signal peptidase. The impairment was even more evident when two charged residues were introduced in the hydrophobic core of the signal sequence. In contrast to deletion mutants (26), this change interferes not only with translocation and cleavage by signal peptidase but also with binding to SRP. Since substitution of a hydrophobic amino acid, leucine, for cysteine or its deletion was ineffective in modifying translocation and processing (26), the present findings would suggest that a change in the hydrophobicity of the core is responsible for the observed global disruption in the processing of the mutant prepro-PTH. Similarly, single charged amino acids introduced in the hydrophobic core of the *E. coli* maltose binding protein signal peptide impair secretion of the protein into the external periplasmic compartment of the cell (27).

It is rather intriguing that three distinct reports of inherited mutations in the signal sequence of human secreted proteins, namely prepro-PTH (9), preprovasopressin (10), and preprofactor X (11), demonstrate inheritance of the associated disorder in an autosomal dominant fashion. As is the case for the other two reported disorders, however, it remains unclear why individuals with a mutated PTH allele have hypoparathyroidism. From transfection studies in COS-7 cells, it would appear that

expression of the mutant allele does lead to secretion of PTH, albeit inefficiently. Moreover, the normal allele would be expected to produce sufficient circulating PTH to maintain calcium homeostasis. The only other reported case of familial isolated hypoparathyroidism segregating with a mutation in the PTH gene, involved a point mutation affecting intron splicing and was associated with autosomal recessive inheritance of the disorder (28). Since heterozygous individuals for this mutant allele were unaffected, it would appear that one normal PTH allele is sufficient for maintaining calcium homeostasis.

Hypoparathyroidism in the presence of one normal PTH allele would therefore suggest that the mutant gene product exerts a dominant negative effect *in vivo*. The mutant protein might interfere with the normal targeting and processing of other secreted proteins, including the normal PTH precursor. Such interference might even lead to destruction of parathyroid tissue in affected individuals; unfortunately, this is difficult to evaluate because the tissue is not readily accessible. Export incompatibility, however, has been observed in *E. coli* expressing transport-defective β -galactosidase leading to lethal jamming of the cellular export machinery (29).

The phenotype of the single-mutant prepro-PTH suggests that it might have dominant negative effects under appropriate conditions. The mutation allows a fraction of the prepro-PTH precursor to enter the translocation machinery, but the mutant protein is then cleaved inefficiently by signal peptidase. The inability of the uncleaved protein to reach the glycosylation machinery (assessed using the precursor modified by inclusion of a glycosylation signal) suggests that the precursor is not transported fully across the microsomal membrane. Such a protein that engages the translocation apparatus but fails to

THIS PAGE BLANK (USPTO)

move through the apparatus efficiently might well have dominant negative effects. The competition experiment in Fig. 9 failed to demonstrate such an effect; perhaps higher levels of protein expression or a longer term experiment are needed. Nevertheless, the observation that all three reported human signal sequence mutations involve proteins that partly engage the secretory apparatus and appear to have dominant effects suggests that *in vivo* these mutant proteins cause dominant secretory dysfunction.

REFERENCES

- Walter, P., Gilmore, R., and Blobel, G. (1984) *Cell* 38, 5-8
- Gilmore, R. (1993) *Cell* 75, 589-592
- Walter, P., and Blobel, G. (1981) *J. Cell Biol.* 91, 557-561
- Connolly, T., and Gilmore, R. (1989) *Cell* 57, 599-610
- Rapiejko, P. J., and Gilmore, R. (1992) *J. Cell Biol.* 117, 493-503
- Miller, J. D., Wilhelm, H., Gierasch, L., Gilmore, R., and Walter, P. (1993) *Nature* 368, 351-354
- Simon, S. M., and Blobel, G. (1991) *Cell* 65, 371-380
- Simon, S. M., and Blobel, G. (1992) *Cell* 69, 677-684
- Arnold, A., Horst, S. A., Gardella, T. J., Baba, H., Levine, M. A., and Kronenberg, H. M. (1990) *J. Clin. Invest.* 86, 1084-1087
- Ito, M., Oiso, Y., Murase, T., Kondo, K., Saito, H., Chinzei, T., Racchi, M., and Lively, M. O. (1993) *J. Clin. Invest.* 91, 2565-2571
- Racchi, M., Watzke, H. H., High, K. A., and Lively, M. O. (1993) *J. Biol. Chem.* 268, 5735-5740
- Ahn, T. G., Antonarakis, S. E., Kronenberg, H. M., Igarashi, T., and Levine, M. (1986) *Medicine* 65, 73-81
- Walter, P., Ibrahimi, I., and Blobel, G. (1981) *J. Cell Biol.* 91, 545-550
- Zoller, M. J., and Smith, M. (1983) *Methods Enzymol.* 100, 468-500
- Baba, H., Karaplis, A. C., Wiren, K. M., Keutmann, H. T., and Kronenberg, H. M. (1992) *J. Bone Miner. Res.* 7, 199-206
- Wiren, K. M., Potts, J. T., Jr., and Kronenberg, H. M. (1988) *J. Biol. Chem.* 263, 19771-19777
- Walter, P., and Blobel, G. (1983) *Methods Enzymol.* 96, 84-93
- Connolly, T., and Gilmore, R. (1986) *J. Cell Biol.* 103, 2253-2261
- Perara, E., and Lingappa, V. R. (1985) *J. Cell Biol.* 101, 2292-2301
- Jackson, R. C., and Blobel, G. (1977) *Proc. Natl. Acad. Sci. U. S. A.* 74, 5598-5602
- Luthman, H., and Magnusson, G. (1983) *Nucleic Acids Res.* 11, 1295-1307
- Lim, S. K., Gardella, T. J., Baba, H., Nussbaum, S. R., and Kronenberg, H. M. (1992) *Endocrinology* 131, 2325-2330
- Meyer, D. I., Krause, E., and Dobberstein, B. (1982) *Nature* 297, 647-650
- Perlman, D., and Halvorson, H. O. (1983) *J. Mol. Biol.* 167, 391-409
- von Heijne, G. (1985) *J. Mol. Biol.* 184, 99-105
- Cioffi, J. A., Allen, K. L., Lively, M. O., and Kemper, B. (1989) *J. Biol. Chem.* 264, 15052-15058
- Bedouelle, H., Bassford, P. J., Jr., Fowler, A. V., Zabin, I., Beckwith, J., and Hofnung, M. (1980) *Nature* 285, 78-81
- Parkinson, D. B., and Thakker, R. V. (1992) *Nature Genet.* 1, 149-152
- Bassford, P. J., Jr., Silhavy, T. J., and Beckwith, J. (1979) *J. Bacteriol.* 139, 19-31

THIS PAGE BLANK (USPTO)

MOLECULAR BIOLOGY

A translation of *Molekulyarnaya Biologiya*

August 30, 1994

Volume 28, Number 2, Part 2

March-April, 1994

CONTENTS

Engl./Russ.

Study of Biogenesis and Secretion of Alkaline Phosphatase and its Mutant Forms in <i>Escherichia coli</i> . II. Effect of Amino Acid Substitutions in the Processing Site and the N Terminus of Mature Polypeptide Chain on its Biogenesis – A. L. Karamyshev, A. E. Kalinin, I. M. Tsfasman, V. N. Ksenzenko, and M. A. Nesmeyanova	245	362
Study of Biogenesis and Secretion of Alkaline Phosphatase and its Mutant Forms in <i>Escherichia coli</i> . III. Substitutions of Alkaline Phosphatase N-Terminal Amino Acid Affect Enzyme Biogenesis – A. L. Karamyshev, A. E. Kalinin, M. I. Khmel'nitskii, M. G. Shlyapnikov, V. N. Ksenzenko, and M. A. Nesmeyanova	253	374
Anion-Exchange Chromatography Induces DNA-Binding Activity of Proteins – G. M. Rott, V. K. Podgorodnichenko, I. V. Rozhinskaya, and I. S. Novikova	259	383
Structural-Functional Organization of Smallpox Virus Genome. 3. Sequencing and Nucleotide Analysis of a Conservative Region of <i>Hind</i> III Fragments F, N, and A of Strain India-1967 – S. N. Shchelkunov, S. M. Resenchuk, A. V. Totmenin, A. A. Kolykhalov, I. V. Frolov, S. M. Dryga, V. V. Volchkov, V. E. Chizhikov, V. V. Gutorov, V. M. Blinov, and L. S. Sandakhchiev	265	392
Short Glucocorticoid-Responsive Element in the 5'-Flanking Region of Murine Metallothionein-I Gene – S. Yu. Plisov, T. I. Merkulova, and A. L. Shkapenko	274	407
Polyelectrolyte Invariance of Acid-Base Equilibrium in Natural DNA with Different GC Content – S. V. Slonitskii and D. B. Boikova	278	413
<i>Bsp</i> Family of Tandemly Repeated DNA Sequences of Baikal Cottoid Fishes (<i>Cottoidea</i>) – S. Ya. Slobodyanyuk, M. E. Pavlova, A. N. Fedorov, and S. I. Belikov	282	419
Expression of Chimeric (Mouse/Human) Immunoglobulin Genes in Lymphoid and Nonlymphoid cells – S. M. Deev, A. Liber, B. V. Radko, and O. L. Polyakovskii	289	429
Phosphinothricin-Resistant Transgenic <i>Nicotiana tabacum</i> and <i>Solanum tuberosum</i> Plants – L. Padegimas, O. A. Shul'ga, and K. G. Skryabin	294	437
Cloning of Universal Hybridization Probes for Testing Mycoplasmal Contamination in Cell Cultures – S. M. Kalachikov, A. E. Kel', G. M. Dymshits, A. M. Ishov, and S. N. Borkhsenius	298	444
Overexpression of <i>Bacillus intermedius</i> 7P Ribonuclease (Binase) in <i>Escherichia coli</i> – A. A. Shul'ga, A. L. Okorokov, K. I. Panov, F. T. Kurbanov, B. K. Chernov, K. G. Skryabin, and M. P. Kirpichnikov	303	453
Human Cholesterol Esterase/Lipase C-Terminal Fragment Containing the Antigenic Determinant: Cloning, Sequencing and Expression in <i>Escherichia coli</i> – M. K. Chekhranova, E. N. Il'ina, E. R. Shuvalova, D. Yu. Pankov, and Yu. A. Pankov	311	464
Cloning of the Gene for <i>Bacillus circulans</i> Extracellular RNase – N. D. Fedorova, A. A. Shul'ga, M. P. Peredel'chuk, L. V. Kozharinova, P. N. Gol'shin, N. F. Ryabchenko, and M. P. Kirpichnikov	315	468
Temperature Dependence of Acid-Base Equilibrium in DNA – S. V. Slonitskii and V. Yu. Kuptsov ...	318	472

The Russian press date (podpisano k pechatii) of this issue was 03/02/1994. Publication therefore did not occur prior to this date, but must be assumed to have taken place reasonably soon thereafter.

THIS PAGE BLANK (USPTO)

Study of Biogenesis and Secretion of Alkaline Phosphatase and its Mutant Forms in *Escherichia coli*. III. Substitutions of Alkaline Phosphatase N-Terminal Amino Acid Affect Enzyme Biogenesis

A. L. Karamyshev,¹ A. E. Kalinin,¹ M. I. Khmel'nitskii,¹ M. G. Shlyapnikov,¹ V. N. Ksenzenko,¹
and M. A. Nesmeyanova¹

UDC 577.21

Translated from *Molekulyarnaya Biologiya*, Vol. 28, No. 2, pp. 374–382, March–April, 1994.

Original article submitted July 4, 1993.

The effect of N-terminal amino acid substitutions in the *Escherichia coli* alkaline phosphatase [EC 3.1.3.1] on its biogenesis was studied. Using *E. coli* strains carrying different *amber* suppressors and the *phoA* gene with an *amber* mutation at position corresponding to the N-terminal amino acid in mature protein, Arg(+1) was changed to Ser, Gln, Tyr, Leu, Gly, Ala, Glu, Phe, His, Cys, Lys, or Pro. All these amino acid substitutions do not prevent secretion and formation of active enzyme. However, introduction of Pro in position +1 completely inhibits signal peptide cleavage (processing) and results in accumulation of the precursor inside the cells. Other amino acid substitutions have no effect on processing. All N-terminal amino acid substitutions studied change the alkaline phosphatase isozyme spectrum. The experimental evidence obtained suggests that the protease splits off the alkaline phosphatase N-terminal arginine is ineffective with Glu and Lys at position +1.

Key words: alkaline phosphatase; *E. coli*; amino acid substitutions; suppressor tRNA; biogenesis

Alkaline phosphatase is a typical secretory protein of *E. coli*. Like most secretory proteins, it is synthesized as a precursor (prePhoA) with an extra N-terminal sequence named signal peptide [1]. Upon translocation of the protein across the cytoplasmic membrane, the signal peptide is off cleaved by a membrane-bound leader peptidase [2]. It has been shown that the signal peptide is essential for secretion, and most amino acid substitutions disturb this process [1, 3, 4]. Significantly less is known about the effect on alkaline phosphatase translocation across the membrane and signal peptide cleavage caused by amino acid substitutions in its mature part: there are only few publications on this problem [5, 6]. We have found that amino acid substitutions within the leader peptidase recognition site and mature alkaline phosphatase N-terminal domain considerably affect distinct steps of its biogenesis, suggesting an important role of these domains [6]. In line with this, the alkaline phosphatase N-terminal amino acid is of great interest.

This is a positively charged arginine relatively rare at this position in secretory proteins. In the process of biogenesis, the N-terminal arginine is removed and active alkaline phosphatase macromolecules are formed through subunit dimerization, yielding three enzyme isoforms [7]. In addition, this amino acid is a part of the signal peptide cleavage site. It has been assumed for a long time that the nature of a secretory protein N-terminal amino acid is not critical for processing [8]. Recently, however, it was found for the phage M13 coat protein [9] and *E. coli* maltose-binding protein [10] that some substitutions at this position interfere with the signal peptide cleavage.

In this work the effect of N-terminal Arg substitutions on particular steps of alkaline phosphatase biogenesis is studied.

MATERIALS AND METHODS

Bacterial strains and plasmids. Construction and characterization of *E. coli* strains used in this work were described in detail previously [11].

¹Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, Pushchino, Moscow Region, 142292.

THIS PAGE BLANK (USPTO)

Plasmid p15KS(+) obtained from Dr. Fisher (FGR) contains multiple cloning sites identical to those of pBluescript KS (Stratagene), p15 *ori* of replication, chloramphenicol resistance marker. Plasmids pPHOA12 and pPHOA7 were constructed on the basis of p15KS(+). They contain the wild-type *phoA* or mutant *phoA7(am)* gene of *E. coli*, respectively [11]. *PhoA7* mutation changes the CGG codon corresponding to alkaline phosphatase Arg(+1) to an *amber* codon TAG. Plasmid pGFIB1 and its derivatives containing *E. coli* *amber* suppressor tRNA genes Ala2, GluA, Phe, Gly1, HisA, ProH, and Cys [12] were obtained from J. Miller (USA). The *amber* suppressor tRNA gene pT5Su of phage T5 was obtained by oligonucleotide-directed mutagenesis and cloned in vector pGFIB1. Amino acid specificities of *amber* suppressor tRNAs used in the study are shown in Table 1. Mutant proteins are designated according to the main amino acid used by *amber* suppressor tRNAs and the position where amino acid substitution was introduced: PhoASer(+1), PhoAGln(+1), PhoALeu(+1), PhoATyr(+1), PhoAAla(+1), PhoAPro(+1), PhoAGly(+1), PhoAGlu(+1), PhoAPhe(+1), PhoACys(+1), PhoAHis(+1), PhoALys(+1).

TABLE 1. *Amber* Suppressors* Used in the Study

<i>Amber</i> suppressor	Amino acid incorporated
Ala2	Ala
GluA	Glu (59%), Gln (17%), Tyr (6%), Arg (6%)
Phe	Phe
Gly1	Gly
HisA	His
ProH	Pro
Cys	Cys
Ser(Su1)	Ser
Gln(Su2)	Gln
Tyr(Su3)	Tyr
Leu(Su6)	Leu
Lys(pT5Su)**	Lys

* Construction of *amber* suppressors was described by Kleina et al. [12]. Amino acid specificity was determined by Normanly et al. [13].

** Phage T5 *amber* suppressor whose amino acid specificity was determined in this work.

Cultivation of *E. coli* in synthetic medium was described in detail in our previous article [6].

Determination of alkaline phosphatase processing rate *in vitro* was done by a published method [4] with modifications. *E. coli* cells grown to mid-exponential phase in synthetic medium [14] were transferred to the same but orthophosphate-free medium and incubated for 10 min at 37°C. This time interval corresponds to the start of alkaline phosphatase synthesis. For protein pulse-labeling, [³⁵S]methionine (Radiopreparat) was added at 40 µCi/ml. After incubation for 30 sec at the same temperature, [³⁵S]methionine incorporation was terminated by adding cold amino acid to 0.05%. At

certain time intervals, 50 µl aliquots were taken, precipitated with 10% trichloroacetic acid, immunoprecipitated [15], and analyzed by electrophoresis followed with autoradiography.

Analytical methods. Protein electrophoresis in 10% polyacrylamide gel with sodium dodecyl sulfate was carried out according to Laemmli [16]. Immunoprecipitation with rabbit anti-alkaline phosphatase antibodies was done as described previously [15]. Immunoblotting was done by electrophoretic transfer of proteins from gel to nitrocellulose BA-85 (Schleicher und Schull) followed with alkaline phosphatase detection, using rabbit antibodies and protein A-horseradish peroxidase conjugate (Bio-Rad Laboratories) as described by Tsfasman et al. [17]. Alkaline phosphatase isoforms were visualized directly in polyacrylamide gel by treatment with α -naphthyl phosphate and Fast Red RR [18] after electrophoresis under native conditions. Alkaline phosphatase activity was determined from the rate of *p*-nitrophenyl phosphate hydrolysis: one unit (U) is the amount of enzyme that hydrolyzes 1 µmole of substrate in 1 min at 37°C. Protein was assayed by the Lowry method [20].

RESULTS

Construction of Mutant Alkaline Phosphatases Using *Amber* Suppressors

To obtain mutant alkaline phosphatases, an approach was used that includes (a) introducing an *amber* mutation into the *phoA* gene by oligonucleotide-directed mutagenesis, and (b) using *amber* suppressors to incorporate a desired amino acid in the corresponding position. The introduction of the *amber* mutation was described previously [11]; its location is shown in Fig. 1a. The mutant *phoA7(am)* gene was cloned in vector p15KS(+) at *Hind*III and *Bam*HI sites (Fig. 1b). The resulting plasmid pPHOA7 has a p15A *ori* of replication and a chloramphenicol resistance marker. The *phoA7(am)* gene with its own promoter P_{phoA} has the same orientation as the P_{lac} promoter. The P_{phoA} promoter enables induction of alkaline phosphatase synthesis in a minimal medium lacking orthophosphate [19]. In suppressor-less *E. coli* strain, the *amber* mutation in the *phoA* gene results in premature translation termination, while mutant proteins are synthesized in strains containing *amber* suppressors. In this work both "classical" *amber* suppressors, whose genes have chromosomal localization (Su1, Su2, Su3, Su6), and *amber* suppressors whose genes are cloned in the plasmid pGFIB1 under the control of constitutive P_{lpp} promoter, from the collection of Kleina and coauthors [12], were used. The plasmid pGFIB1 bears an ampicillin resistance marker and the ColE1 *ori* of replication allowing its use in a two-plasmid system in combination with pPHOA7 (Fig. 1b). The amino acid specificities of *amber* suppressor tRNAs were established by Normanly et al. [13] and

THIS PAGE BLANK (USPTO)

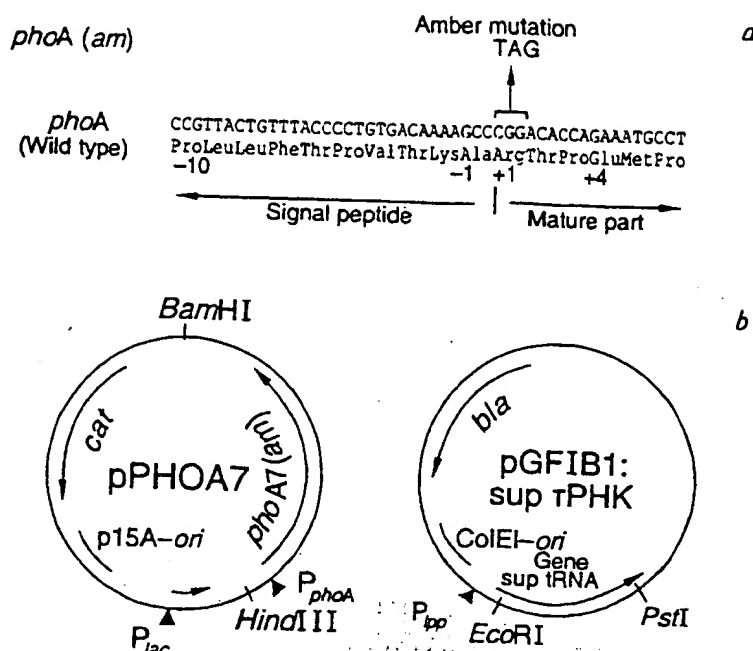


Fig. 1. a) Location of *amber* codon in the *phoA* gene. The structure of the wild-type *PhoA* gene 5'-terminal region corresponding to amino acids Pro(-10)—Pro(+6) is presented; nucleotide changes resulting in *amber* codon appearance are shown. b) Two-plasmid system used for production of mutant alkaline phosphatases.

are presented in Table 1. We have described the construction of *E. coli* cells with the *phoA* gene deleted containing respective *amber* suppressors and pPHOA7 [11]. Using the approach described above, the alkaline phosphatase N-terminal Arg was replaced with Ser, Gln, Tyr, Leu, Gly, Ala, Glu, Phe, His, Cys, Lys, and Pro.

Common Features in Secretion and Biogenesis of Mutant Alkaline Phosphatases Obtained Using Suppressor tRNAs

All mutant proteins studied have enzymic activity. This conclusion follows from the phosphatase activity assays in bacterial culture and culture liquid of *E. coli* strains XAC (MK376-MK379 and MK391-MK398) harboring the plasmid pPHOA7 (Table 2). Phosphatase activity in *E. coli* strains MC1061/p3 (MK372-MK375 and MK381-388) and E15 (MK401-MK408) differs considerably from the data presented above. This, probably, is associated with the different efficiency of amino acid incorporation by *amber* suppressor tRNAs in different *E. coli* strains. Though this makes impossible a quantitative comparison of expression and secretion for different mutant proteins, their enzymic activities suggest that none of the amino acid substitutions obtained results in complete block of alkaline phosphatase translocation across the cytoplasmic membrane and active enzyme formation.

It has been shown that the wild-type *phoA* gene expression from multicopy plasmids results in enhanced alkaline phosphatase synthesis and alters its biogenesis: the *PhoA* precursor (prePhoA) is accumulated inside the cells as insoluble aggregates, and mature protein is

secreted not only to periplasm but to culture liquid as well (up to 80% of active enzyme) [14, 21]. As can be seen in Table 2, the efficiency of both wild-type and mutant alkaline phosphatase secretion to the medium in the strains studied in this work is significantly lower than in other *E. coli* strains studied previously [21]. This may be associated with the peculiarities of cell wall structure in the strains used. It was also shown by immunoblotting

TABLE 2. Activities of Mutant Alkaline Phosphatases in Bacterial Culture and Culture Liquid*

Amber suppressor	Alkaline phosphatase activity, mU/mg protein	
	Culture	Medium
Su1 (Ser)	1319	426
Su2 (Gln)	2798	94
Su3 (Tyr)	1703	110
Su6 (Leu)	5034	150
Ala2	4563	556
GluA	2534	359
Phe	1740	206
Gly1	1509	265
HisA	2290	317
ProH	1913	94
Cys	1259	111
Lys(pT5Su)	3605	380
Control	5573	280

*Assayed in *E. coli* strains MK376-MK379 and MK391-MK399 harboring the plasmid pPHOA7. *E. coli* strain MK380 containing the plasmid with the wild-type *phoA* gene (pPHOA2) was used as a control.

THIS PAGE BLANK (USPTO)

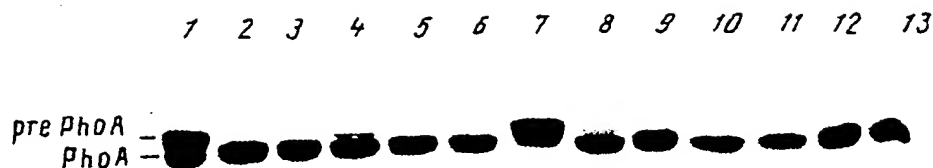


Fig. 2. Immunoblotting of the wild-type and mutant alkaline phosphatases in cell lysates. 1) Wild-type PhoA; 2) PhoASer(+1); 3) PhoAGln(+1); 4) PhoALeu(+1); 5) PhoATyr(+1); 6) PhoAAla(+1); 7) PhoAPro(+1); 8) PhoAGly(+1); 9) PhoAGlu(+1); 10) PhoAPhe(+1); 11) PhoACys(+1); 12) PhoAHis(+1); 13) PhoALys(+1).

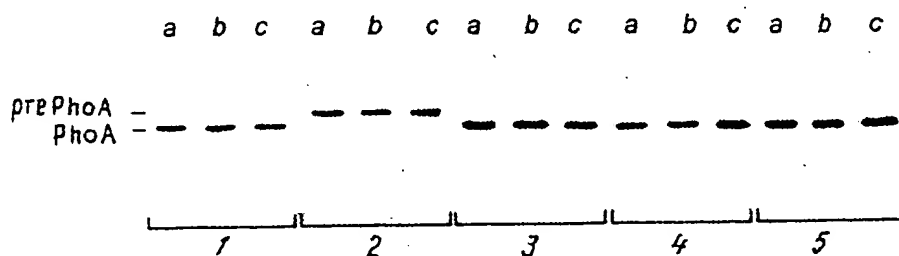


Fig. 3. Analysis of mature alkaline phosphatase formation from ^{35}S -labeled precursors of: 1) wild-type PhoA; 2) PhoAPro(+1); 3) PhoAAla(+1); 4) PhoAGln(+1); 5) PhoALeu(+1) in 0 (a), 0.5 (b), and 60 min (c) after termination of labeling with $[^{35}\text{S}]$ methionine *in vivo*. Samples were analyzed by electrophoresis in 10% polyacrylamide gel followed with autoradiography.

that upon suppression of the *amber* mutation most of the mutant alkaline phosphatase precursors are not accumulated inside the cells, or their amounts are negligible, whereas the wild-type precursor is abundant (Fig. 2). This can be explained by the fall in the protein synthesis rate upon the *amber* suppressor tRNA-aided amino acid incorporation, resulting in complete translocation of the synthesized protein across the cytoplasmic membrane.

Mutant PrePhoA Processing *in vivo*

Translocation of precursor across the cytoplasmic membrane and its conversion to mature form through signal peptide cleavage are important steps of secretory protein biogenesis. To study the effect of amino acid substitutions on these processes, *in vivo* pulse-labeling of protein was carried out. *E. coli* cells transformed with respective plasmids were labeled with $[^{35}\text{S}]$ methionine for 30 sec and then incubated in the absence of the label. After immunoprecipitation, the alkaline phosphatase precursor and mature form were separated electrophoretically. As can be seen in Fig. 3, labeled PhoAPro(+1) is detected as precursor only, not converted to the mature form throughout the time of observation, suggesting the lack of signal peptide cleavage. This conclusion is also supported by the absence of this protein mature form when the cell proteins are analyzed by immunoblotting (Fig. 2). The rates of precursor processing for the wild-type alkaline phosphatase and proteins PhoAAla(+1), PhoAGln(+1), PhoALeu(+1) (Fig. 3) as well as other mutant proteins

with Arg(+1) changed to Ser, Tyr, Gly, Glu, Phe, His, Cys, or Lys (data not shown) were very high, and their precursors were not revealed under these conditions. These data suggest that no of the amino acid substitutions studied except Pro is critical for the *in vivo* protein processing.

Posttranslational Modification of Mutant Alkaline Phosphatases

The last step in the wild-type alkaline phosphatase biogenesis includes removal of the N-terminal Arg and dimerization of protein subunits, yielding three enzyme isoforms: dimer with Arg at the N termini of both subunits (isoform I), heterodimer of subunits with and without N-terminal Arg (isoform II), and dimer in which both subunits are devoid of N-terminal Arg (isoform III) [7]. To study the isozyme spectra of mutant alkaline phosphatases, we separated periplasmic proteins by electrophoresis under nondenaturing conditions and detected the active alkaline phosphatase directly in the gel (Fig. 4). The wild-type PhoA is present in the periplasm as three isoforms; the mutant isozyme spectra differ considerably from that of the wild-type protein. The PhoAGlu(+1) is revealed as a major isoform with a higher electrophoretic mobility than that of the wild-type isoform III. The presence of a minor isoform just above the major one can be explained by the amino acid specificity of *amber* suppressor tRNA^{GluA}, which can accept glutamine (17%) in addition to glutamic acid (59%) [13]. The PhoALys(+1) is detected as a single form with the same electrophoretic mobility as the

THIS PAGE BLANK (USPTO)

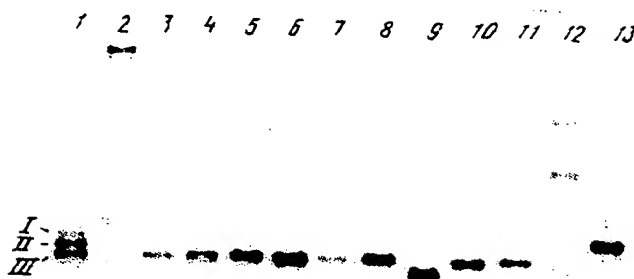


Fig. 4. Comparison of mutant and wild-type alkaline phosphatase isozyme spectra. 1) Wild type PhoA; 2) PhoAPro(+1); 3) PhoASer(+1); 4) PhoAGln(+1); 5) PhoATyr(+1); 6) PhoALeu(+1); 7) PhoAGly(+1); 8) PhoAAla(+1); 9) PhoAGlu(+1); 10) PhoAPhe(+1); 11) PhoAHis(+1); 12) PhoACys(+1); 13) PhoALys(+1). The wild-type alkaline phosphatase isoforms are denoted by Roman numerals. Samples were analyzed by electrophoresis under native conditions followed with alkaline phosphatase visualization by treatment with α -naphthyl phosphate and Fast Red RR.

wild-type isoform I. Analysis of the isozyme spectra of mutant alkaline phosphatases with charged amino acids (Glu and Lys) substituted for Arg(+1) suggests that N-terminal amino acid removal by the modifying protease Iap is impaired, otherwise other isoforms would have appeared in the case of N-terminal Lys or Glu. It should be noted that an *E. coli* strain containing the Lys *amber* suppressor pT5Su was used in this work. The gene of this suppressor tRNA was obtained by oligonucleotide-directed mutagenesis of the phage T5 tRNA^{Lys} gene. The amino acid specificity of this suppressor tRNA was not established before. The isozyme spectrum analysis indicates that this suppressor tRNA does accept Lys, otherwise a different isozyme spectrum would have been observed. In the case of noncharged amino acids or histidine, the mutant enzyme should comigrate with the wild-type PhoA isoform III, in the case of Glu or Asp it should be below isoform III, and in the case of Arg all three isoforms of the wild-type protein would be detected. According to Normanly et al. [13], the *amber* suppressor tRNA obtained from *E. coli* tRNA^{Lys} also does not change its amino acid specificity.

Isozyme spectra of proteins containing Ser, Gln, Tyr, Ala, Phe, or Gly at position +1 look as expected for proteins with substitution of a noncharged amino acid for Arg(+1). All of them are detected as single isoforms with the same electrophoretic mobility as the wild-type isoform III. The same electrophoretic pattern is found for PhoAHis(+1). Since the pK' for histidine R-group is 6.0, it is uncharged under the electrophoretic conditions (pH 8.3), and hence does not affect the protein electrophoretic mobility. In contrast with the mutant proteins mentioned above, introducing Pro or Cys in position +1 significantly alters the isozyme spectrum: several isoforms of these proteins are detected. In the case of PhoAPro(+1), this can be explained by protein self-aggregation or aggregation with other cell

components owing to the presence of uncleaved signal peptide. In the case of PhoACys(+1), perhaps formation of additional disulfide bonds between several subunits of this mutant protein takes place. The formation of such bonds, however, does not result in the loss of enzymic activity. In this case, the N-terminal methionine also is not removed by the modifying protease.

DISCUSSION

After crossing the cytoplasmic membrane, secretory protein precursors undergo processing that results in signal peptide cleavage and conversion of the protein to the mature form. This reaction is catalyzed by a leader (signal) peptidase [2, 22]. In our previous work [6] the high structural conservedness of amino acids at positions -3 and -1 within the signal peptidase recognition site was discussed. Position +1 is currently believed to be much less conserved [23, 24]. Indeed, all 20 amino acids are met at this position in eukaryotic secretory protein precursors [8]. However, Gln, Ile, Leu, Met, and Pro are not found at position +1 of prokaryotic secretory proteins [8, 25]. Our data suggest that a change of Arg(+1) to Ala, Glu, Phe, His, Cys, Ser, Gln, Leu, or Tyr in alkaline phosphatase does not significantly affect the protein processing *in vivo*. However, proline in this position completely inhibits the signal peptide cleavage. This finding is in full agreement with the data obtained in studying the processing of phage M13 coat protein [9] or maltose-binding protein [10]. Moreover, as found in the latter study, the mechanism of processing inhibition for the maltose-binding protein with proline at position +1 is different from that for the protein with amino acid substitutions at positions -3, -1 of the signal peptide. While the latter mutations interfere with recognition of the signal peptide cleavage site, the presence

THIS PAGE BLANK (USPTO)

of Pro(+1) makes the mutant protein uncleavable, but does not impair its recognition by signal peptidase. Such a mutant protein acts as a competitive inhibitor of leader peptidase, which leads to accumulation of several secretory protein precursors inside the cells. The mechanism of PhoAPro(+1) processing inhibition is likely to be similar to the one described above; however, special experiments are required to clarify this question.

In pulse-labeling experiments, secretory protein precursors are detected if either the protein translocation across the cytoplasmic membrane or the signal peptide cleavage by leader peptidase is blocked [9]. In this work it was found that inhibition of processing of the Arg(+1)-Pro protein does not prohibit the precursor translocation across the cytoplasmic membrane, as witnessed by the high phosphatase activity of this protein in bacterial cells. The *E. coli* alkaline phosphatase is known to become active only in the periplasm, where conditions are appropriate for disulfide bonding and active macromolecule formation, while the nontranslocated alkaline phosphatase precursor has no enzymic activity [5, 7]. Moreover, the nontranslocated precursor is usually degraded by cytoplasmic proteases in 5–10 min [4]. As can be seen in Fig. 3, the mutant PhoAPro(+1) precursor does not undergo proteolysis even after 60 min of incubation, suggesting once more its translocation across the cytoplasmic membrane. Our data agree with the scanty findings for other proteins in which the introduction of proline in position +1 inhibits processing but not translocation [9, 10].

After completion of processing, alkaline phosphatase undergoes further proteolytic modification: N-terminal arginine removal by protease Iap [7]. It has been shown that adding arginine to the culture medium inhibits the conversion of alkaline phosphatase isoform I to isoforms II and III: from this observation the specificity of protease Iap was suggested [26, 27]. However, the specificity of Iap protease toward N-terminal arginine has not been directly proved until now. Our data suggest that at least glutamic acid or lysine introduced in position +1 of alkaline phosphatase are not removed by the modifying protease.

The data of this work indicate that all substitutions of N-terminal arginine in alkaline phosphatase are not absolutely critical for its translocation across the cytoplasmic membrane and for active molecule formation; however, they affect certain steps of biogenesis.

In this study, amino acids were incorporated into the position corresponding to *amber* mutation in the *phoA* gene with the aid of *amber* suppressor tRNAs. Using a set of *amber* suppressors, we were able to obtain a broad set of proteins with Arg(+1) substituted with various amino acids. Despite the great number of *amber* suppressors available [12, 28, 29] and the obvious attractiveness of their use in protein engineering, this approach has not been widely applied heretofore. Our data show that introduction of amino acid changes with the aid of different *amber* suppressor tRNAs is promising

when a quantitative evaluation of mutant protein expression is not necessary. Moreover, the system used by us to obtain the mutant alkaline phosphatases can be applied for assessing the specificity of unidentified *amber* suppressors by analysis of the mutant alkaline phosphatase N-terminal amino acids. This is technically simpler than determination of the 10th amino acid in the *E. coli* dihydrofolate reductase, whose gene with *amber* mutation in this position is currently used for such purpose [13].

The study was done with financial support from the Russian Fund for Basic Research.

REFERENCES

1. L. M. Gierash, *Biochemistry*, 28, 923–930 (1989).
2. T. Date and W. Wickner, *Proc. Natl. Acad. Sci. USA*, 78, 6106–6110 (1981).
3. D. Oliver, *Annu. Rev. Microbiol.*, 39, 615–648 (1985).
4. S. Michaelis, J. F. Hunt, and J. Beckwith, *J. Bacteriol.*, 167, 160–167 (1986).
5. P. Li, J. Beckwith, and H. Inouye, *Proc. Natl. Acad. Sci. USA*, 85, 7685–7689 (1988).
6. A. L. Karamyshev, A. E. Kalinin, I. M. Tsfasman, V. N. Ksenzenko, and M. A. Nesmeyanova, *Mol. Biol.*, 28, 364–365 (1994).
7. A. Toriani-Gorini, F. G. Rothman, S. Silver, A. Wright, and E. Yagil (editors), *Phosphate Metabolism and Cellular Regulation in Microorganisms*, American Society for Microbiology, Washington (1987).
8. M. E. E. Watson, *Nucl. Acids Res.*, 12, 5145–5164 (1984).
9. L. M. Shen, J.-I. Lee, S. Cheng, H. Jutte, A. Kuhn, and R. E. Dalbey, *Biochemistry*, 30, 11775–11781 (1991).
10. G. A. Barkocy-Gallagher and P. J. Bassford, Jr., *J. Biol. Chem.*, 267, 1231–1238 (1992).
11. A. L. Karamyshev, M. G. Shlyapnikov, M. I. Khmel'nitskii, M. A. Nesmeyanova, and V. N. Ksenzenko, *Mol. Biol.*, 28, 150–157 (1994).
12. L. G. Kleina, J.-M. Masson, J. Normanly, J. Abelson, and J. H. Miller, *J. Mol. Biol.*, 213, 705–717 (1990).
13. J. Normanly, L. G. Kleina, J.-M. Masson, J. Abelson, and J. H. Miller, *J. Mol. Biol.*, 213, 719–726 (1990).
14. M. A. Nesmeyanova, I. M. Tsfasman, A. L. Karamyshev, and N. E. Suzina, *World J. Microbiol. Biotechnol.*, 7, 394–406 (1991).
15. K. Ito, P. Bassford, Jr., and J. Beckwith, *Cell*, 24, 707–714 (1981).
16. U. K. Laemmli, *Nature*, 227, 680–685 (1970).
17. I. M. Tsfasman, M. A. Nesmeyanova, V. G. Gorbulev, P. M. Rubtsov, and K. G. Skryabin, *Mol. Biol.*, 23, 422–430 (1989).
18. Z. Loida, R. Gossrau, and T. Schibler, *Enzyme Histochemistry* [Russian translation], Moscow, Mir (1982).
19. A. Toriani, *Biochim. Biophys. Acta*, 38, 460–466 (1960).
20. O. H. Lowry, N. J. Rosebrough, A. L. Farr, and R. J. Randall, *J. Biol. Chem.*, 193, 265–275 (1951).
21. M. A. Nesmeyanova, A. L. Karamyshev, I. M. Tsfasman, and E. T. Fedotikova, *Mol. Biol.*, 25, 770–778 (1991).
22. R. E. Dalbey and G. von Heijne, *TIBS*, 17, 474–478 (1992).
23. G. von Heijne, *Eur. J. Biochem.*, 133, 17–21 (1983).
24. G. von Heijne, *J. Mol. Biol.*, 173, 243–251 (1984).
25. G. von Heijne, *Nucl. Acids Res.*, 14, 4683–4690 (1986).
26. P. J. Piggot, M. D. Sklar, and L. Gorini, *J. Bacteriol.*, 110, 291–299 (1972).
27. M. J. Schlessinger, W. Bloch, and P. M. Kelly, In: C. L. Markert (editor), *Isozymes. I. Molecular Structure*, New York, Academic Press (1975), pp. 333–342.
28. J. H. Miller, J. M. Masson, L. G. Kleina, J. Normanly, and J. Abelson, *J. Cell. Biochem., Suppl.* 11C, 195–203 (1987).
29. J. M. Masson, L. G. Kleina, J. H. Miller, J. Normanly, and J. N. Abelson, *Protein Eng.*, 1, 233–242 (1987).

THIS PAGE BLANK (USPTO)

Signal Peptidases Recognize a Structural Feature at the Cleavage Site of Secretory Proteins*

(Received for publication, August 10, 1987, and in revised form, March 30, 1988)

Guy Duffaud† and Masayori Inouye

From the Department of Biochemistry, Robert Wood Johnson Medical School at Rutgers, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey 08854-5635

The cloning of the gene for staphylococcal nuclease A in the pIN-III-OmpA secretion vector results in a hybrid protein which is processed by signal peptidase I, yielding an active form of the nuclease that is secreted across the cytoplasmic membrane (Takahara, M., Hibler, D., Barr, P. J., Gerlt, J. A., and Inouye, M. (1985) *J. Biol. Chem.* 260, 2670-2674). Using oligonucleotide-directed site-specific mutagenesis, we have constructed a set of mutants at the cleavage site area of the precursor hybrid protein designed to alter progressively the predicted secondary structure of the cleavage site. Our results show that processing becomes increasingly defective as the turn probability decreases. These results are consistent with the structural requirement that we found for the processing of lipoprotein by signal peptidase II (Inouye, S., Duffaud, G., and Inouye, M. (1986) *J. Biol. Chem.* 261, 10970-10975). We conclude that secretory precursor proteins have a distinct secondary structural requirement at their cleavage site for processing by signal peptidase I, as well as by signal peptidase II.

Secretory proteins are usually synthesized as precursors containing an extra amino-terminal extension termed the signal peptide (Watson, 1984). During protein secretion across the cytoplasmic membrane of *Escherichia coli*, specific proteases are required for the removal of these peptide extensions. These proteases are called signal peptidases; and in *E. coli*, two have been shown to be responsible for processing of all secretory precursors: signal peptidase I (leader peptidase) (Wolfe *et al.*, 1982) and signal peptidase II (Husain *et al.*, 1980). All secretory proteins but lipoproteins are processed by signal peptidase I. Lipoproteins are processed by signal peptidase II, which requires a lipid modification at the cleavage site of the precursor before processing (reviewed by Wu (1987)).

Prokaryotic signal peptides consist of approximately 20-30 amino acid residues and do not show a significant sequence homology (reviewed by Duffaud *et al.* (1985)). However, they are organized in three discrete regions: 1) an amino-terminal region containing 1-3 positively charged amino acid residues; 2) a hydrophobic sequence of 14-20 residues often punctuated by 1 or 2 proline and/or glycine residues; and 3) a cleavage site region usually joined to the hydrophobic region by a serine and/or threonine residue. The carboxyl-terminal end

of the signal peptide (cleavage site) is always an amino acid with a small side chain: serine, alanine, or glycine. This feature at the cleavage site has been shown to be essential for the processing of the precursor of the major outer membrane lipoprotein of *E. coli* (Pollitt *et al.*, 1986). Only two amino acids could replace the glycine at the carboxyl terminus of the signal peptide: alanine and serine. Substitution by threonine, valine, or leucine resulted in defective processing. Unlike the C terminus of the signal peptide, there is no amino acid favored at the N terminus of the mature portion of the secretory protein (Duffaud *et al.*, 1985). An additional structural feature has been described for the cleavage site: a high probability for a turn structure (Inouye *et al.*, 1986). We have demonstrated that as the predicted turn probability at the cleavage site decreases, there is a hampering in the processing of the precursor of the major lipoprotein of *E. coli*. When the predicted turn structure was completely eliminated, the mutant precursor protein became totally defective in lipid modification and processing.

In this study, we attempt to examine whether signal peptidase I recognizes similar structural features at the cleavage site. For this purpose, we created several mutations that progressively diminished the probability for the formation of a turn structure at the cleavage site of the OmpA signal peptide, which is known to be processed by signal peptidase I. Analysis of these mutants showed that a similar structural feature at the cleavage site as required for signal peptidase II also exists at the cleavage site region of precursor proteins that are processed by signal peptidase I.

MATERIALS AND METHODS

Strain, Plasmid, and Medium—*E. coli* strain SB221 (*lpp hsdR⁻ AtrpE5 leuB6 lacY recA1/F⁺ lacI^s lac⁺ pro⁺*) (Nakamura *et al.*, 1982) was used as host for all plasmids in this study. Cells were grown as previously described (Inouye *et al.*, 1982). Plasmid pIN-III-OmpA-#98 contains the nuclease A gene of *Staphylococcus aureus* fused with a linker to the OmpA signal peptide. The resulting protein is synthesized as a precursor which is processed by signal peptidase I and yields an active form of the nuclease which is exported to the periplasmic space of *E. coli* (Takahara *et al.*, 1985). Plasmids derived by mutagenesis are described according to the following systematic rules: 1) numbers always refer to the wild-type amino acid position and follow the described alteration; 2) substitution is indicated by the single-letter code name of the new amino acid; and 3) deletion is symbolized by the Greek capital delta (Δ) (Duffaud *et al.*, 1985). In addition, plasmids derived from pIN-III-OmpA-#98 are named according to the localization of the mutation: A if it is in the OmpA signal peptide and N if it is in the mature portion of the hybrid protein. For example, one of the plasmids in this study has the proline residue at position 24 deleted and accordingly is termed pAN(Δ 24) (see Fig. 1) (Lehnhardt *et al.*, 1987).

Oligonucleotide-directed Site-specific Mutagenesis—Oligonucleotides were synthesized on a Systec Microsyn-1450-DNA synthesizer by the phosphoramidate method and purified by preparative gel electrophoresis as previously described (Sinha *et al.*, 1984; Inouye and

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported by a Setsuro Fujii fellowship. Present address: Dept. of Biochemistry, Brandeis University, Waltham, MA 02154-9110.



THIS PAGE PLANK #1(SPT0)

Expression of Mutant Genes and Analysis of Protein Products—Pulse-chase labeling experiments were carried out as described previously (Lehnhardt *et al.*, 1987). Immunoprecipitation was done using goat antinucleic sera, and the results were analyzed by SDS-PAGE¹ using a 17% polyacrylamide gel (Inouye *et al.*, 1982).

RESULTS

The gene coding for the OmpA-nuclease hybrid was constructed by cloning the nuclease gene into a high expression secretion vector, pIN-III-*ompA*3 (Takahara *et al.*, 1985). The hybrid protein produced by pIN-III-*OmpA*-#98 consists of the signal peptide of the OmpA protein, a major outer membrane protein, fused to nuclease A of *S. aureus*. It has been shown that the OmpA signal peptide is accurately processed by signal peptidase I and that the resulting product is secreted into the periplasmic space. The secreted product contains 11 extra amino acid residues at the N terminus of nuclease A, resulting from the linker sequence between the OmpA signal sequence and the nuclease A. This peptide extension is underlined in Fig. 1. This hybrid protein is referred to here as "wild type," as opposed to the mutants derived from it.

Expression of OmpA-Nuclease Mutants—Since the gene for the hybrid protein is under the control of both the *lpp* promoter and the *lac* promoter-operator, the expression of this gene can be induced in *E. coli* SB221 by adding isopropyl-1-thio- β -D-galactopyranoside (IPTG), a *lac* inducer, to the cul-

+15

424

Δ24Δ25

K22Δ24Δ25

K22

PheAlaThrValAlaGlnAla **Lys**IleProSerLeuAspProThrValTyrSerAlaThr
CGCAGGCC AAGATTCCA

ture medium. All the mutants were found to produce active nuclease when induced. However, one mutant was found to be lethal: *E. coli* SB221 carrying pAN(K22Δ24Δ25) was not

¹ The abbreviations used are: SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; IPTG, isopropyl- β -D-thiogalactopyranoside.

THIS PAGE BLANK (USPTO)

capable of forming colonies above 30 °C on L-broth plates containing IPTG at 2 mM. The other mutants were able to form colonies at all three temperatures tested (30, 37, and 42 °C). This result suggested that the product of pAN(K22Δ24Δ25), albeit an active nuclease, was toxic to the host, probably by accumulating the enzyme inside the cell and/or by accumulating the unprocessed precursor in the membrane.

Pulse-Chase Experiment—A pulse-chase experiment of the wild-type hybrid and of mutants Δ24, Δ24Δ25, and K22Δ24Δ25 was performed at 37 °C, as shown in Fig. 3. Cells harboring the wild-type plasmid (#98) or one of its mutant derivatives were grown to a klett reading of 50 and were then induced for 10 min with IPTG. The cells were then pulse-labeled with [³⁵S]methionine for 30 s. After the pulse labeling, the cells were chased with nonradioactive methionine for 30 s, 60 s, 90 s, 2 min, and 5 min. The reactions were stopped by addition of cold trichloroacetic acid, and the nuclease proteins were immunoprecipitated and analyzed by SDS-PAGE.

It is clear from Fig. 3 that processing of the mutant proteins is slower in all cases than processing of the wild-type protein, with the effect being particularly marked for mutant K22Δ24Δ25, which even after 5 min, has less than 50% of its precursor protein processed to the mature form (lane 24, arrowhead b). Furthermore, it is evident that the rate of processing seems to be reduced for each new alteration added to the wild-type protein, as can be seen when comparing lanes 3 (wild type), 9 (one deletion), 15 (two deletions), and 21 (two deletions and one substitution). We determined the half-life of each precursor protein by calculating the ratio of precursor to mature form in each lane after densitometric measurement of each band. The estimated half-lives are found to be as follows: 30 s, #98; 50 s, Δ24; 120 s, Δ24Δ25; and >5 min, K22Δ24Δ25. This progressive decrease in the rate of processing can be correlated with the decrease in the probability for a turn structure at the cleavage site, as shown in Fig. 2.

The dramatic effect observed in the processing rate of mutant K22Δ24Δ25 may probably explain its lethal effect at 37 and 42 °C. It is likely that the mutant hybrid protein can still be translocated across the cytoplasmic membrane but not processed, remaining attached to the membrane. Accumulation of such unprocessed precursor molecules in the cytoplasmic membrane has been shown to be lethal to the cell (Inouye *et al.*, 1983).

It is possible that the alterations introduced in the primary structure of the hybrid protein could have affected the processing of the mutant protein by changing the location of the processing site. To address this possibility, we performed an

Edman protein degradation of the labeled and processed products of the wild-type plasmid (pAN) and the pAN(K22Δ24Δ25) mutant labeled with [³H]isoleucine. In both sequenced proteins, a distinct peak appeared at the second cycle, indicating that both the wild-type plasmid and the pAN(K22Δ24Δ25) mutant are processed at the expected cleavage site (Table I).

Effect of Mutant pAN(K22)—The dramatic effect brought by the substitution of a glycine at position 22 by a lysine in mutant Δ24Δ25 led us to investigate the role of this particular position. It is quite possible that replacing the glycine with a lysine residue might make the cleavage site less recognizable by signal peptidase I; for example, a bulky side chain at that position could "hide" the site. In order to examine this possibility, we constructed mutant K22. The individual structural contribution of K22 can be estimated by the rules of Chou and Fasman (1978) as done above. A lysine at position 22 contributes to the overall decrease in turn probability close to the cleavage site. However, the probability for an α-helix after the cleavage site does not increase as much as for Δ24 (data not shown). Overall, the structural effect of K22, if any, might be comparable to the effect of Δ24. Thus, K22 might have a potential for affecting processing in two ways: an accessibility factor for the peptidase and a structural recognition effect.

The result of a pulse-chase experiment on mutant K22 is shown in Fig. 4. It is clear that mutation K22 by itself does not alter significantly the rate of processing (compare lanes 1 and 2 with 7 and 8, respectively). Thus, the effect of K22 can be exerted only when it is combined with Δ24Δ25. This result also suggests that the nature of the amino acid at the N terminus of the cleavage site is not essential for recognition by signal peptidase I of the cleavage. However, it should be

TABLE I

Edman degradation of the pAN(K22Δ24Δ25) product

A 5-ml volume of cell culture at mid-log phase was induced with IPTG for 3 min. At that point, 200 μl of [³H]isoleucine were added (32 Ci/mmol, 1 mCi/ml). After a 25-min incubation, the cells were treated as described in the text, and the plasmid protein products were immunoprecipitated and electroblotted from an SDS-PAGE experiment to a polyvinylidene difluoride membrane. The membrane was cut, and the absorbed proteins were manually sequenced by the Edman method (details in text). The values represent the average 10⁴ min counts after background counts were subtracted.

Cycle	pIN-III-OmpA-#98	pAN(K22Δ24Δ25)
First	22	3
Second	235	102
Third	11	0
Fourth	24	32

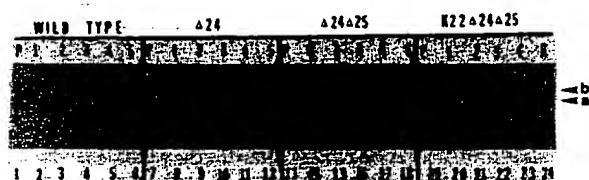


FIG. 3. Immunoprecipitates from pulse-chase experiments at 37 °C of mutants Δ24, Δ24Δ25, and K22Δ24Δ25. *E. coli* SB221 harboring pIN-III-OmpA-#98 or the mutant plasmid was pulse-labeled with [³⁵S]methionine and chased at different intervals of time, as described in the text. The induction time was 10 min. The pulse time (lane P) was 30 s followed by a chase at 30 s (upper lane 1), 60 s (upper lane 2), 90 s (upper lane 3), 2 min (upper lane 4), and 5 min (upper lane 5). Lanes 1–6 refer to the wild-type hybrid, lanes 7–12 to mutant Δ24, lanes 13–18 to Δ24Δ25, and lanes 19–24 to K22Δ24Δ25. The mature and precursor proteins are indicated by arrowheads a and b, respectively.

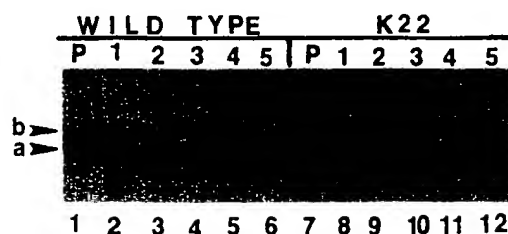


FIG. 4. Immunoprecipitates from pulse-chase experiment at 37 °C of mutant K22. *E. coli* SB221 harboring pIN-III-OmpA-#98 or the mutant plasmid was pulse-labeled with [³⁵S]methionine and chased at different intervals of time, as described in the text. The induction time was 10 min. The pulse time (lane P) was 30 s, followed by a chase at 30 s (upper lane 1), 60 s (upper lane 2), 90 s (upper lane 3), 2 min (upper lane 4), and 5 min (upper lane 5). Lanes 1–6 refer to the wild-type hybrid and lanes 7–12 to mutant K22. The mature and precursor proteins are indicated by arrowheads a and b, respectively.



THIS PAGE BLANK (USPTO)

TABLE II

Recognition requirements for processing by signal peptidases

The data presented in this report are summarized below along with the results already known about the requirements of signal peptidases as discussed throughout the text.

	Signal peptidase	
	I	II
Structural requirement	Turn	Turn
Primary structural requirement		
C terminus of signal peptide	Gly, Ala, ^a Ser	Gly, Ala (Ser) ^b
NH ₂ terminus of mature portion	Not apparent ^c	Cys
Specific sequence	Not apparent	Leu-Leu-Ala-Gly↓Cys ^d
Other	Not known	Lipid modification at N terminus Cys

^a The most frequently found amino acid residues are underlined.

^b Serine has not been found to occur naturally in lipoproteins, but was shown to be tolerated when substituted at that position by mutagenesis.

^c Possibly dependent on satisfying the structural requirement.

^d This sequence is referred to as the *lipoprotein box*. It is partially conserved among the lipoproteins, and it is thought to be important for the lipid modification enzymes and signal peptidase II. The arrow indicates the cleavage site.

pointed out that the N-terminal amino acid of the mature secretory protein can be altered only as long as the overall structural requirement remains satisfied.

DISCUSSION

Processing of secretory precursor proteins by signal peptidases takes place at a precise site, although there is no consensus sequence in the primary structure of the signal peptide. In this regard, signal peptidases are unique among peptidases and proteases, which generally require specific primary sequences to recognize their cleavage sites. The question is then how signal peptidases are able to recognize the processing site of the signal peptide. A secondary structural requirement for a turn structure was proposed in the case of the lipoprotein which is processed by signal peptidase II (Inouye *et al.*, 1986). In this work, it was shown that the probability for a turn at the cleavage site could be decreased by amino acid substitution in the mature portion of the protein, and this resulted in a reduced rate in the lipid modification and processing of the lipoprotein precursor. Based on these facts, it is possible to think that a similar structural requirement exists for the recognition by signal peptidase I. Actually, examination of several precursor proteins according to the rules of Chou and Fasman (1978) showed a high probability for a turn structure at the cleavage site of proteins processed by signal peptidase I; an example is given in Fig. 2A for the hybrid protein OmpA-staphylococcal nuclease A, which is processed by the same signal peptidase. In order to test if a turn structure is also required for signal peptidase I, three mutations have been created on the mature portion of the OmpA-nuclease hybrid. The mutations were designed to alter progressively the predicted turn structure at the cleavage site, as shown in Fig. 2 (B-D) for mutants pAN(Δ 24), pAN(Δ 24 Δ 25), and pAN(K22 Δ 24 Δ 25), respectively. The most severe effect is observed in the processing rate of the pAN(K22 Δ 24 Δ 25) mutant, which had a lethal effect on the cell when expressed at 37 or 42 °C (see Fig. 3). It should be emphasized that the severe effect observed for mutant K22 Δ 24 Δ 25 is not due to the substitution of glycine at position 22 by a lysine. This mutation by itself does not have an effect, as seen in Fig. 4. The K22 mutation, as a "silent" mutation, becomes lethal only when combined with another mutation that might not be lethal by itself either, as is the case with mutant Δ 24 Δ 25. The results presented here are consistent with the prediction that signal peptidase I also

has a turn structure requirement at the cleavage region of the precursor protein, as is the case for signal peptidase II (Inouye *et al.*, 1986). Thus, the alterations introduced in the OmpA-nuclease protein affected processing of the precursor by disrupting its secondary structure without changing the actual cleavage site.

A structural requirement for a turn at the cleavage site for both peptidases is also consistent with the proposed loop model for protein secretion (Inouye *et al.*, 1979; Inouye and Haleboua, 1980). In this model, the signal peptide forms a loop during its insertion in the membrane at the cleavage site. It is conceivable that the loop structure at the cleavage site is exposed to the outside of the cytoplasmic membrane, where signal peptidase I was shown to be located (Wolfe *et al.*, 1982). The loop could well generate a turn structure that could be recognized by a signal peptidase.

From the results presented here and other known properties of signal peptidases I and II also discussed here, it is possible to summarize their requirements, as shown on Table II. Signal peptidase I requires a turn structure at the cleavage site, in addition to a small side chain amino acid residue (Gly, Ala, and Ser) at the N terminus of the signal peptide. Signal peptidase II, in addition to the same requirements as signal peptidase I, requires a specific primary sequence at the cleavage site region. This unique sequence is required for the lipid modification, which is in turn required for the processing by signal peptidase II (reviewed by Wu (1987)). It is to be expected that the mature portion of a precursor protein will play a role in the processing and secretion of the protein. Our results are also comparable with previous reports. Bankaitis *et al.* (1984) have isolated pseudorevertants located in the mature portion of the protein that compensate for secretion-defective signal peptide mutations. Interestingly, the same signal peptide does not necessarily have the same effect if it is attached to two different proteins (Lehnhardt *et al.*, 1987). We have reported the importance of a compatibility between the signal peptide and the secretory protein to generate the appropriate structure required for the processing by a signal peptidase (Inouye *et al.*, 1986). We have also shown that the amino acid residue at the carboxyl terminus of the lipoprotein signal peptide can be Gly, Ser, or Ala, but not Thr, Val, or Leu (Pollitt *et al.*, 1986). The carboxyl terminal alanine of the OmpA signal peptide can also be replaced by a glycine residue without affecting the processing rate of the precursor.² From

² G. Duffaud and M. Inouye, unpublished observations.

THIS PAGE BLANK (USPTO)

these results, signal peptidases can be described as a new class of proteases which primarily recognize the secondary structure of a protein, namely a reverse turn structure. The cleavage reaction will proceed if there is also a Gly, Ala, or Ser residue appropriately located in the specific turn formed by the looping precursor protein.

REFERENCES

- Bankaitis, V. A., Rasmussen, B. A., and Bassford, P. J. (1984) *Cell* **37**, 243-252
- Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251-276
- Duffaud, G. D., Lehnhardt, S. K., March, P. E., and Inouye, M. (1985) *Curr. Top. Membr. Transp.* **24**, 65-104
- Husain, M., Ichihara, S., and Mizushima, S. (1980) *J. Biol. Chem.* **255**, 3707-3712
- Inouye, M., and Halegoua, S. (1980) *CRC Crit. Rev. Biochem.* **7**, 339-371
- Inouye, S., and Inouye, M. (1987) in *Synthesis and Application of DNA and RNA* (Narang, S., ed) pp. 181-204, Academic Press, New York
- Inouye, M., Pirtle, I., Sekizawa, J., Nakamura, K., Di Rienzo, J., Inouye, S., Wang, S., and Halegoua, S. (1979) *Microbiology Wash. D.C.* **34**-37
- Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K., and Inouye, M. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79**, 3438-3441
- Inouye, S., Franceschini, T., Sato, M., Itakura, K., and Inouye, M. (1983) *EMBO J.* **2**, 87-91
- Inouye, S., Duffaud, G., and Inouye, M. (1986) *J. Biol. Chem.* **261**, 10970-10975
- Konigsberg, W. (1967) *Methods Enzymol.* **11**, 461-469
- Lehnhardt, S., Pollitt, S., and Inouye, M. (1987) *J. Biol. Chem.* **262**, 1716-1719
- Matsudaira, P. (1987) *J. Biol. Chem.* **262**, 10035-10038
- Nakamura, K., Masui, Y., and Inouye, M. (1982) *J. Mol. Appl. Genet.* **1**, 289-299
- Pollitt, S., Inouye, S., and Inouye, M. (1986) *J. Biol. Chem.* **261**, 1835-1837
- Sanger, F., Nicklen, S., and Coulson, S. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463-5467
- Sinha, N. D., Biernat, J., and Koster, H. (1984) *Nucleic Acids Res.* **12**, 4539-4557
- Smith, J. A., and Pease, L. G. (1980) *CRC Crit. Rev. Biochem.* **8**, 315-399
- Takahara, M., Hibler, D., Barr, P. J., Gerlt, J. A., and Inouye, M. (1985) *J. Biol. Chem.* **260**, 2670-2674
- Watson, M. E. E. (1984) *Nucleic Acids Res.* **12**, 5145-5164
- Wu, H. C. (1987) in *Bacterial Outer Membrane as Model Systems* (Inouye M., ed) pp. 37-71, Wiley-Interscience, New York
- Wolfe, P. B., Silver, P., and Wickner, W. (1982) *J. Biol. Chem.* **257**, 7898-7902

THIS PAGE BLANK (USPTO)

13. Zlotnik, I. & Stamp, J. T. *Vet. Rec.* 77, 1178-1179 (1965).
14. Zlotnik, I. & Stamp, J. T. *Vet. Rec.* 78, 222 (1966).
15. Burger, D. & Gorham, J. R. *Res. vet. Sci.* 22, 131-132 (1977).
16. Gajdusek, D. C., Gibbs, C. J. Jr., Collins, G. & Traub, R. D. *New Engl. J. Med.* 294, 553 (1976).
17. Bernoulli, C. et al. *Lancet* i, 478-479 (1977).
18. Gard, S. & Aaloe, O. in *The Viruses* Vol. 1 (eds Burnet, F. M. & Stanley, W. M.) 359-427 (Academic, New York, 1959).
19. Kim, K. S. & Sharp, D. G. *J. Immun.* 99, 1221-1225 (1967).
20. Burger, D., Gorham, J. R. & Leader, R. W. in *Slow, Latent and Temperate Virus Infections* (eds Gajdusek, D. C., Gibbs, C. J. Jr & Alpers, M.) 307-313 (NINDB Monogr. 2, US Government Printing Office, Washington DC, 1965).
21. Salk, J. E. & Gori, J. B. *Ann. N.Y. Acad. Sci.* 83, 609-637 (1960).
22. Rohwer, R. G. *Science* 223, 600-602 (1984).
23. Alper, T., Haig, D. A. & Clarke, M. C. *J. gen. Virol.* 41, 503-516 (1978).
24. Field, E. J., Farmer, F., Caspary, E. A. & Joyce, G. *Nature* 222, 90-91 (1969).
25. Lea, D. E. *Actions of Radiations on Living Cells* (Cambridge University Press, 1946).
26. Ginoza, W. in *Methods in Virology* Vol. 4 (eds Maramorosch, K. & Koprowski, H.) 139-209 (Academic, New York, 1968).
27. Kaplan, H. S. & Moses, L. E. *Science* 145, 21-25 (1964).
28. Terzi, M. *Nature* 191, 461-463 (1961).
29. Rohwer, R. G. & Gajdusek, D. C. *Proc. 1st int. Symp. Herie Foundation* (ed. Boese, A.) 333-355 (Verlag Chemie, Weinheim, 1980).
30. Tischer, I., Gelderblom, H., Vettermann, W. & Koch, M. A. *Nature* 295, 64-66 (1982).
31. Griffith, I. S. *Nature* 215, 1043-1044 (1967).
32. Pattison, I. H. & Jones, K. M. *Vet. Rec.* 80, 2-9 (1967).
33. Diring, H., Kimberlin, R. H. *Biosci. Rev.* 3, 563-568 (1983).
34. Malone, T. G., Marsh, R. F., Hanson, R. P. & Semancik, J. S. *J. Virol.* 25, 933-935 (1978).
35. Prusiner, S. B. et al. *Biochemistry* 17, 4993-4999 (1978).
36. Kim, K. S. & Sharp, D. G. *Radiat. Res.* 33, 30-36 (1968).
37. Kim, K. S. & Sharp, D. G. *J. Virol.* 1, 45-49 (1967).
38. Prusiner, S. B. et al. *Proc. natn. Acad. Sci. U.S.A.* 78, 4606-4610 (1981).
39. Millson, G. C. & Manning, J. E. in *Slow Transmissible Diseases of the Nervous System* Vol. 2 (eds Prusiner, S. B. & Hadlow, W. J.) 409-424 (Academic, New York, 1979).
40. Prusiner, S. B. et al. *J. Neurochem.* 35, 574-582 (1980).
41. Kimberlin, R. H., Millson, G. C. & Hunter, G. D. *J. comp. Path.* 81, 383-391 (1971).
42. Rohwer, R. G., Brown, P. W. & Gajdusek, D. C. in *Slow Transmissible Diseases of the Nervous System* Vol. 2 (eds Prusiner, S. B. & Hadlow, W. J.) 465-477 (Academic, New York, 1979).
43. Eklund, C. M., Hadlow, W. J. & Kennedy, R. C. *Proc. Soc. exp. Biol.* 112, 974-979 (1963).
44. Gibbs, C. J. Jr, Gajdusek, D. C. & Morris, J. A. in *Slow, Latent and Temperate Virus Infections* (eds Gajdusek, D. C., Gibbs, C. J. Jr & Alpers, M.) 195-202 (NINDB Monogr. 2, US Government Printing Office, Washington DC, 1965).
45. Wilson, D. R., Anderson, R. D. & Smith, W. J. *comp. Path. Ther.* 60, 267-282 (1950).
46. Merz, P. A., Somerville, R. A., Wisniewski, H. M. & Iqbal, K. *Acta neuropath.* 54, 63-74 (1981).
47. Merz, P. A., Somerville, R. & Wisniewski, H. M. in *Virus Non Conventions et Affections du System Nervous Central* (eds Court, L. A. & Cathala, F.) 259-281 (Masson, Paris, 1983).
48. Merz, P. A., Somerville, R. A., Wisniewski, H. M., Manuelidis, L. & Manuelidis, E. E. *Nature* 306, 474-476 (1983).
49. Merz, P. A. et al. *Science* (in the press).
50. Diring, H., Gelderblom, H., Hilmert, H., Ozel, M. & Edelbluth, C. *Nature* 306, 476-478 (1983).
51. Marsh, R. F. & Kimberlin, R. H. *J. infect. Dis.* 131, 104-110 (1975).
52. Kimberlin, R. H. & Walker, C. A. *J. gen. Virol.* 34, 295-304 (1977).
53. Reed, L. J. & Muench, H. *Am. J. Hyg.* 27, 493-497 (1938).
54. Ferguson, J. & Davis, R. W. *J. molec. Biol.* 94, 135-149 (1975).
55. Delendi, V., Jensen, F. & Saur, G. in *The Molecular Biology of Viruses* Vol. 6 (eds Colter, J. S. & Paranchy, W.) 645-662 (Academic, New York, 1967).
56. Reddy, V. B. et al. *Science* 200, 494-502 (1978).
57. Fraenkel-Conrat, H. *Compreh. Virol.* 1 (1974).
58. Drasil, V., Doskar, J., Viklicka, S. & Tkadlec, L. *Acta virol.* 22, 443-450 (1978).
59. Van der Eb, A. J., Van Kesteren, L. W. & Van Bruggen, E. F. J. *Biochim. biophys. Acta* 182, 530-541 (1969).
60. Freifelder, D. *Virology* 30, 328-332 (1966).
61. Roizman, B. & Furlong, D. *Compreh. Virol.* 3, 229-403 (1974).
62. Becker, Y. & Sarov, I. *J. molec. Biol.* 34, 655-660 (1968).
63. Ginoza, W. *Nature* 199, 453-456 (1963).
64. Diener, T. O. & Schneider, I. *Virology* 29, 100-105 (1966).
65. Hightower, L. E. & Brutt, M. A. *Compreh. Virol.* 9, 535-598 (1977).

Secretion of a wheat α -amylase expressed in yeast

Steven J. Rothstein*, Colin M. Lazarus*,
Wendy E. Smith, David C. Baulcombe
& Anthony A. Gatenby

Plant Breeding Institute, Trumpington, Cambridge CB2 2LO, UK

The translocation of secretory proteins across the endoplasmic reticulum involves the recognition and cleavage of an amino-terminal extension called the signal sequence¹. The structure of signal peptides appears to be ubiquitous in having a very hydrophobic central core², so that the signal sequence in secretory proteins from one organism could possibly be recognized by the processing and transport apparatus of another. We therefore wished to investigate whether a protein, α -amylase, one of several hydrolytic enzymes secreted from the aleurone of wheat into the endosperm during germination, could be processed and secreted in an active form from the yeast *Saccharomyces cerevisiae*, secretion being dependent upon the plant signal sequence. Here, synthesis of α -amylase was by inserting a cDNA clone coding for the entire α -amylase structural gene³ into a yeast expression vector⁴. The α -amylase protein coded for by this gene fusion has the signal sequence located internally, not at the N-terminal end of the polypeptide. Nevertheless, it is processed and the processed form is secreted into the medium in an active form. There are potential industrial applications for yeast that secrete a functional α -amylase.

A previously isolated cDNA clone for a wheat α -amylase gene was inserted into a yeast expression vector as shown in Fig. 1a. The yeast expression vector pMA230 codes for the phosphoglycerate kinase (PGK) promoter as well as the first 12 amino acids of the PGK polypeptide⁴. The α -amylase cDNA was constructed using the dG/dC homopolymer tailing method and cloned into the *Pst*I site of pBR322⁵. The plasmid 2128 contains the entire protein-coding sequence as well as 5' and 3' untranslated regions. The insertion of the cDNA coding for the

α -amylase gene into the single *Bam*HI site of pMA230, making the plasmid 520 (Fig. 1a), creates a hybrid gene coding for a translational fusion product as shown in Fig. 1b. From the sequence of the α -amylase gene (C.M.L., unpublished results), the polypeptide synthesized will have a total of 38 extra amino acids at its N-terminal end, coded for by the PGK gene, the normally untranslated α -amylase leader (in which there are no nonsense codons), the dG/dC homopolymer tail and DNA linker fragments (see Fig. 1b). The α -amylase signal peptide is therefore not at the N-terminal end of the polypeptide.

Another plasmid (528) was constructed in which the dG/dC tails, the 5' untranslated region and almost all of the signal peptide coding region were deleted (Fig. 1a). As shown in Fig. 1b, the polypeptide coded for by 528 should be 23 amino acids longer than the wild-type mature wheat α -amylase. Plasmid 521 (Fig. 1a) is the same as plasmid 520 except that the α -amylase gene was inserted in the wrong orientation to be transcribed by the PGK promoter. Yeast cells carrying this plasmid were used as non- α -amylase producing strains.

Yeast cells transformed with the plasmid 520 were labelled with ³⁵S-methionine, the intracellular protein was isolated, immunoprecipitated with anti- α -amylase serum and electrophoresed on a SDS-polyacrylamide gel. Two distinct polypeptide bands can be seen on the resulting autoradiograph shown in Fig. 2a. The larger polypeptide has an apparent molecular weight (MW) of 49,000 which is the expected value for the predicted fusion protein coded for by 520. The smaller polypeptide is the processed form of the 49,000-MW protein (S.J.R., unpublished results) and is almost identical in molecular weight to mature *in vivo* labelled wheat α -amylase (42,000 MW), although the mobilities of the two differ slightly (Fig. 2b, tracks A, B). This small difference may be caused by differences in other post-translational events such as methylation or glycosylation. When cells transformed with the plasmid 528 were labelled with ³⁵S-methionine a single polypeptide of the expected molecular weight (44,400) was immunoprecipitated by anti- α -amylase serum. As expected since the signal peptide is absent, no cleavage to a smaller protein is seen (Fig. 2c). The incorporation of label into α -amylase in cells carrying plasmid 528 is considerably greater than that seen for cells containing plasmid 520. To demonstrate that this is not due to a general increase in incorporation of the label into protein, the β subunit of the yeast mitochondrial ATPase was co-immunoprecipitated with

Fig. 1
analyse
a, A P
2128 (r
 α -amyla
vector p
There ar
with a s
gene (C
A partial
lated from
single B
4). Plasm
correct c
and correl
lation of
the yeas
(PGK) g
gene is i
PGK pro
tion of 5
the dG/d
untransla
signal pe
RsaI site
the *Pst*I
 α -amylas
unpublish
fragment
digested
three Ha
~10 bp
HaeIII-H
long. Thi
with *Sma*
and after
was isol
HaeIII-H
pUC9. Th
plasmid (
interest w
*Eco*RI (r
*Bam*HI-
a triple li
giving 52
(ref. 8) w
H. HaeII
cleavage s
amino aci
of pMA23
includes t
the wheat
identical (

*Present addresses: Ciba-Geigy Biotechnology, Research Triangle Park, North Carolina 27709, USA (S.J.R.); Botany Department, University of Bristol, Bristol BS8 1UG, UK (C.M.L.).

THIS PAGE BLANK (USPTO)

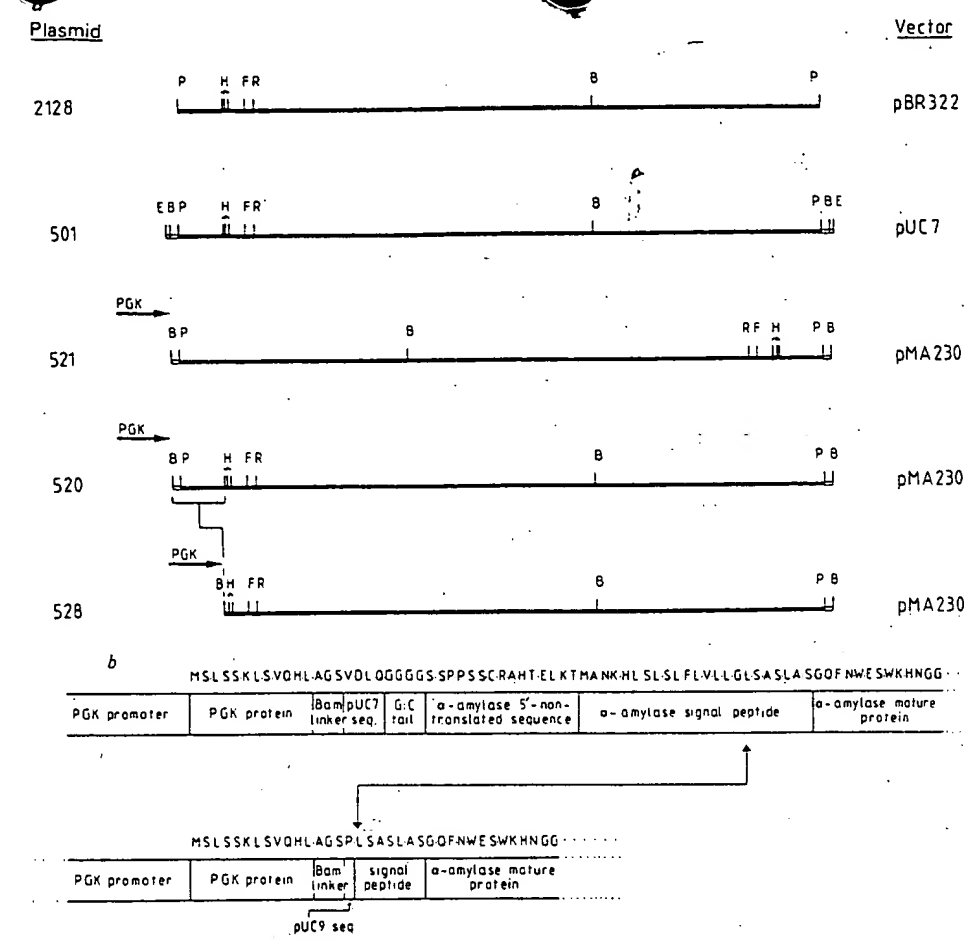
es of the
ic, New
974-979
te V
Mo
1950)
4, 63-74
fections
n, Paris,
is, E. E.
176-478
s Colter,
78).
rys. Acta

aking
for a
n the
sults),
amino
e, the
at
D
ptide is

G/dC
signal
in Fig.
acids
plasmid
at the
to be
ig this

labelled
olated,
elec-
poly-
shown
molecular
or the
ypep-
S.J.R.,
weight
MW),
tracks
nces in
cosyla-
labelled
pected
anti- α -
absent,
20
528
plasmid
ncrease
t of the
ed with

Fig. 1 Plasmids constructed to analyse α -amylase synthesis in yeast. **a**, A *Pst*I fragment from plasmid 2128 (ref. 3) coding for the wheat α -amylase gene was ligated into the vector pUC7 (ref. 7) to make 501. There are two *Bam*HI sites in pUC7, with a single site in the α -amylase gene (C.M.L., unpublished results). A partial *Bam*HI fragment was isolated from 501 and inserted into the single *Bam*HI site of pMA230 (ref. 4). Plasmid 520 is an insertion in the correct orientation for transcription and correct reading frame for translation of the α -amylase gene by the yeast phosphoglycerate kinase (PGK) gene, while the α -amylase gene is inverted with respect to the PGK promoter in 521. The construction of 528 involved the deletion of the dG/dC homopolymer tail, the 5' untranslated leader and most of the signal peptide as follows: there is an *Rsa*I site 302 base pairs (bp) from the *Pst*I site at the 5' end of the α -amylase sequence in 501 (C.M.L., unpublished results). This *Pst*I-*Rsa*I fragment was isolated and partially digested with *Hae*III. There are three *Hae*III sites in this fragment ~10 bp apart, with the largest *Hae*III-*Rsa*I fragment being 75 bp long. This partial digest was ligated with *Sma*I-cleaved pUC9 (ref. 7) and after transformation, a plasmid was isolated, having the largest *Hae*III-*Rsa*I fragment inserted into pUC9. The restriction map of this plasmid (525) at the sequence of interest was found to be the following: *Bam*HI (from pUC9)-*Sma*I/*Hae*III (junction site)-*Hin*II (α -amylase)-*Rsa*I/*Sma*I (junction site)-*Eco*RI (pUC9). The *Hin*II site is 55 bp from the *Hae*III site. The reconstruction of the entire α -amylase gene involved isolating the *Bam*HI-*Hin*II fragment from 525 and the *Hin*II-*Eco*RI fragment from 501 and ligating the two with *Bam*HI plus *Eco*RI-cleaved pUC9 in a triple ligation to make 526. The partial *Bam*HI fragment coding for the α -amylase gene was isolated from 526, and inserted into pMA230 giving 528, which has the α -amylase gene in the correct reading frame to be translationally fused with the PGK protein. The strain MC1022 (ref. 8) was used for the transformation involving pUC7 and pUC9. MC1061 (ref. 8) was used for all other transformations. Symbols: P, *Pst*I; H, *Hae*III; F, *Hin*II; R, *Rsa*I; B, *Bam*HI; E, *Eco*RI; \rightarrow , direction of transcription of the PGK promoter. Note that the only *Hae*III and *Rsa*I cleavage sites shown are those used in the construction. **b**, Predicted α -amylase translation products coded for by plasmids 520 and 528. The amino acid sequence at the N-terminal end of the α -amylase polypeptides coded for by 520 and 528 are derived from the DNA sequences of pMA230 (ref. 4), various linker fragments (refs 4, 7) and of the α -amylase gene from 2128 (C.M.L., unpublished results). The latter sequence includes the dG/dC homopolymer tails and the 5' untranslated region as well as the normal protein coding sequence. The site of cleavage of the wheat α -amylase signal peptide is predicted from the known cleavage site of the barley enzyme whose DNA sequence in this region is identical (P. Chandler, personal communication).



α -amylase as an internal control (Fig. 2b, c). The cytoplasmic α -amylase activity in cells containing plasmid 528 is about 10 times higher than when 520 is present (data not shown). No detectable α -amylase was precipitated in cells transformed with plasmid 521.

The cleavage of the polypeptide coded for by 520 to a size similar in molecular weight to the wild-type wheat protein raised the possibility that the yeast cell was also secreting the α -amylase. To examine this, yeast cells transformed with 520 were labelled with 35 S-methionine and then incubated with non-selective medium (YPDA) containing unlabelled methionine. The protein in the medium and the intracellular protein were separately reacted with anti- α -amylase serum. As Fig. 2d shows, a radioactive polypeptide of the same molecular weight as the processed α -amylase was immunoprecipitated from the medium. Since the intracellular cytoplasmic fraction possessed both unprocessed and processed forms (as well as a third polypeptide of intermediate molecular weight that was occasionally seen), while the extracellular medium fraction possessed only the processed polypeptide, the latter must have been secreted from the cell rather than appearing due to cell leakage or rupture. One interesting phenomenon is that the secretion of

the labelled polypeptides into the medium appears to be more efficient when the chase is carried out in rich medium rather than selective YGM medium.

We needed to investigate whether the secreted α -amylase was functionally active. Yeast cells containing the plasmids 520, 521 and 528 were therefore streaked on YPGA plates containing 1% starch. The plates were reacted with iodine vapour which stains starch purple. As Fig. 3 shows, yeast colonies containing 520 had clear halos around them indicating degradation of the starch. There was also an occasional purple colony due to curing of the plasmid and subsequent failure to produce α -amylase. (Such colonies can arise because the plasmid was not maintained by selection on the rich medium.) In the case of yeast transformed with 528 and 521 there was no clearing, although when 528 was present the starch underneath the colony and the colonies themselves turned purple much more slowly. This is presumably due either to slight leakage of the α -amylase out of the cells or to cell rupture. There is 10 times as much α -amylase in cells carrying 528 than in those bearing 520, even though the clearing zones are far larger for the latter. These results show that the wheat α -amylase is secreted in a functionally active form. No noticeable halo was seen around cells

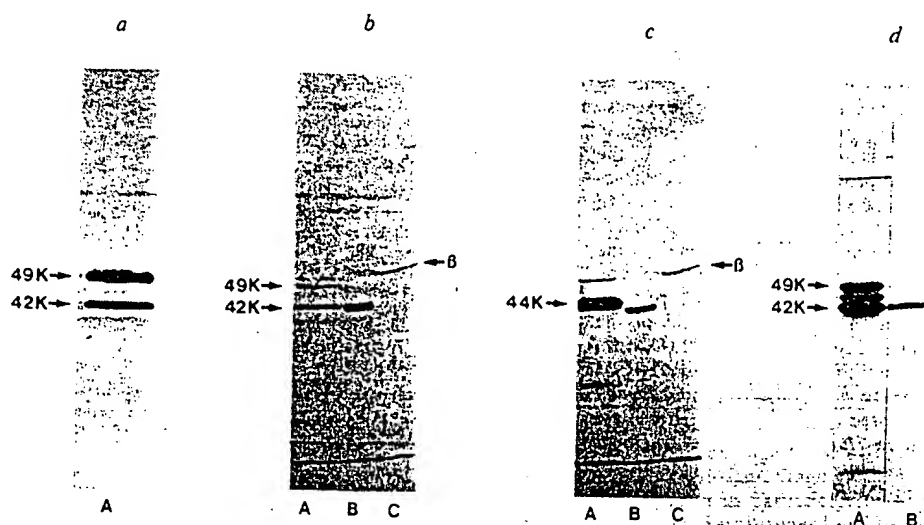
THIS PAGE BLANK (USPTO)

containing 520 plated on selective medium (YGM). This agrees with less α -amylase polypeptide being immunoprecipitated from YGM medium than from a rich medium (see above).

There are two main points about the secretion of the wheat α -amylase in yeast. First, the wheat α -amylase signal peptide is recognized by the yeast-processing apparatus even though the signal peptide is not at the N-terminal end of the polypeptide. This process is slow, however (S.J.R. *et al.*, unpublished results), presumably because the signal peptide is somewhat buried. The time taken to chase half of the labelled precursor to the mature

form was about 30–60 min. The second point is that the processed protein is secreted into the medium in an active form. The amount of enzyme secreted was variable and within the range of 30–60% of total activity. These results are likely to be of interest to the brewing industry since secretion of α -amylase from yeast cells is potentially useful in light beer production⁵. Yeast cells have already been shown to secrete animal secretory proteins⁶. The processing and export of a plant protein points to the general similarity of the eukaryotic secretory signals and processing apparatus.

Fig. 2 α -amylase polypeptides synthesized in yeast transformed with plasmids 520 and 528. The yeast cells were labelled with 35 S-methionine, the cells lysed, the protein immunoprecipitated with anti- α -amylase serum raised in rabbits against the purified wheat enzyme, and electrophoresed on SDS-polyacrylamide gels. **a**, 520 transformed cells labelled for 30 min with 35 S-methionine. **b**, Track A—520 transformed cells labelled for 30 min and then chased with unlabelled methionine for 90 min, the protein being immunoprecipitated with anti-serum against α -amylase and the β subunit of the mitochondrial ATPase; track B, *in vivo* labelled wheat α -amylase immunoprecipitated with anti- α -amylase serum; track C, labelled protein from yeast immunoprecipitated with only the anti- β -serum. **c**, Track A, 528 transformed yeast labelled for 30 min and then chased for 90 min, the protein being immunoprecipitated with both anti- α -amylase and anti- β -serum; track B, *in vivo* labelled wheat α -amylase; track C, labelled yeast protein precipitated only with anti- β -serum. **d**, Track A, cytoplasmic α -amylase protein from 520 transformed yeast labelled for 30 min with 35 S-methionine and then incubated with non-selective (YPDA) medium containing unlabelled methionine for 90 min; track B, the growth medium protein immunoprecipitated with anti- α -amylase serum. Numbers indicate MW ($\times 10^{-3}$).



Methods: The yeast strain used for all these experiments was Mc16 (α *leu* 2-3 *his* 4-712 *ade* 2-1 *lys* 2-1 *SUF2*). The yeast media used were either YPDA (a non-selective, medium containing 1% yeast extract, 2% bacto-peptone, 2% glucose and 20 μ g ml⁻¹ adenine) or YGM (selective medium containing 0.67% yeast nitrogen base w/o amino acids, 2% glucose and 20 μ g ml⁻¹ adenine, histidine and lysine). Yeast protoplasts were transformed essentially as described by Hinnen *et al.*⁹. Cells were grown to $A_{600} = 0.5$ in YGM and then labelled with 10–20 μ Ci 35 S-methionine in a volume of 200 μ l. The intracellular protein was isolated by making protoplasts from the labelled yeast cells with zymolyase 60,000 in an 80- μ l volume for 30 min and then lysing them by adding 20 μ l 10% SDS and boiling for 3 min. After the addition of 1 ml of 1% Triton and 0.15 M NaCl, the insoluble material was removed by centrifugation for 15 min in an Eppendorf microfuge. Immunoprecipitations were carried out as previously described¹⁰. The proteins were electrophoresed on 10% SDS-polyacrylamide gels¹¹, fluoresced using the PPO-DMSO method¹² and dried for autoradiography. *In vivo* labelled wheat α -amylase was prepared by first incubating embryo-less seeds for 36 h in 20 mM sodium acetate (pH 5.5), 20 mM CaCl₂, 0.2% chloramphenicol and 2 μ M gibberellic acid. The aleurone layer was scraped away from the endosperm and 10 aleurones were incubated for 5 h with 100 μ Ci 35 S-methionine. The aleurone protein was isolated by grinding the tissue with sand in 2% SDS and then boiling for 5 min. After removing the debris, by centrifugation the α -amylase was immunoprecipitated as above.

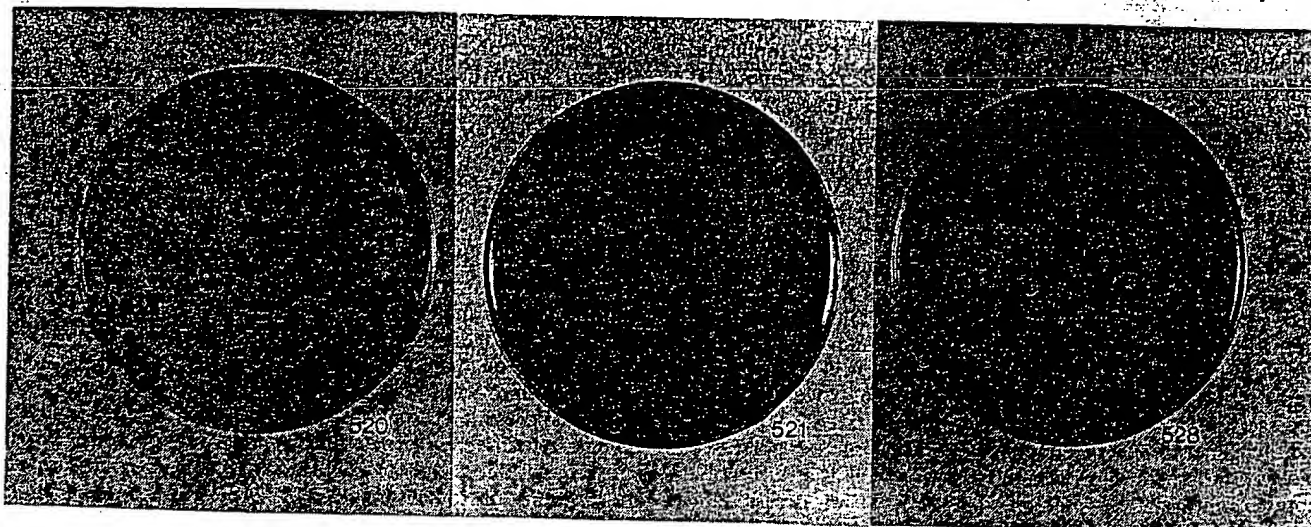


Fig. 3 Yeasts transformed with plasmids 520, 521 and 528 were streaked onto YPDA agar containing 1% soluble starch and 1 mM CaCl₂. After colony growth, the plates were inverted onto beakers containing solid iodine, sitting in a 50 °C water bath. Clear patches indicate starch degradation, while the remaining starch turns a deep purple.

We thank
man for tl
and Dick
script. Thi
a Shell Po

Received 16 N

1. Blobel, G.
2. Finkelstein
3. Baulcomb
4. Tuite, M. I
5. Tubb, R. S
6. Hitzeman
7. Vieira, J. A
8. Casadaban
9. Hinnen, A.
10. Gatenby, A
11. Lacmili, U
12. Bonner, W

Sequen
antigen
in Tryp
do not

Richard G
Murray S
& Nina A

* Departme
Seattle, Was
† Chiron Co
94608, USA

Trypanosom
mechanism
of 'safe' int
or antigenic
somes (for
is mediated
glycoprotein
the active g
same 5' tern
this leader
reiterated t
this array is
functions as
restricting t
rearrangem
and structur
RNA splicin
those which
golense and
cruzi and *Le*
homologous
suggest that
is an ancest
trypanosom

We have
several spec
the *T. brucei*
enzyme-dige
complement
probe hybrid
T. brucei sui
the closely r
1.4 (or 1.5)
directly and
the probe de
(770 base pa
(640 bp), w
(8) it hybridiz

THIS PAGE BLANK (USPTO)

the pro-
ve form.
ith the
el the
-amyase
duction⁵
ecretory
in points
nals and

We thank M. Tuite, M. Dobson, S. Kingsman and A. Kingsman for the yeast expression vector pMA230 used in this work, and Dick Flavell for helpful discussion and reading of the manuscript. This work was in part funded by the NRDC. C.M.L. was a Shell Postdoctoral Fellow.

Received 16 November 1983; accepted 9 February 1984.

1. Blobel, G. & Dobberstein, B. *J. Cell Biol.* 67, 835-851 (1975).
2. Finkelstein, A. V. *et al. FEBS Lett.* 161, 177-179 (1983).
3. Baucombe, D. C. & Buffard, D. *Planta* 157, 493-501 (1983).
4. Tuite, M. F. *et al. EMBO J.* 1, 603-608 (1982).
5. Tubb, R. S. *Critical Reviews in Biotechnology* (CRC Press, Boca Raton, 1983).
6. Hitzeman, R. A. *et al. Science* 219, 620-625 (1983).
7. Vieira, J. & Messing, J. *Gene* 19, 259-268 (1982).
8. Casadaban, M. & Cohen, S. N. *J. molec. Biol.* 138, 179-207 (1980).
9. Hinne, A. *et al. Proc. natn. Acad. Sci. U.S.A.* 75, 1929-1933 (1978).
10. Gatenby, A. A. & Castleton, J. A. *Molec. gen. Genet.* 185, 424-429 (1982).
11. Laemmli, U. K. *Nature* 227, 680-685 (1970).
12. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* 46, 83-88 (1974).

Sequences homologous to variant antigen mRNA spliced leader in Trypanosomatidae which do not undergo antigenic variation

Richard G. Nelson*, Marilyn Parsons*, Murray Selkirk*, George Newport*, Philip J. Barr† & Nina Agabian*

* Department of Biochemistry SJ-70, University of Washington, Seattle, Washington 98195, USA

† Chiron Corporation, 4560 Horton Street, Emeryville, California 94608, USA

Trypanosomes which parasitize mammals have evolved mechanisms to evade immune attack, such as the occupation of 'safe' intracellular sites (for example, *Trypanosoma cruzi*), or antigenic variation, exemplified by the salivarian trypanosomes (for example, *Trypanosoma brucei*). Antigenic variation is mediated by sequential expression of single variant surface glycoprotein (VSG) genes, and often involves transposition of the active gene¹⁻⁶. Every VSG transcript examined shares the same 5' terminal 35-nucleotide leader sequence^{7,8}. In *T. brucei*, this leader is encoded within a 1.4-kilobase unit tandemly reiterated to form a large array^{9,10}. It is hypothesized¹⁰ that this array is distantly linked to the expressed VSG gene and functions as a multiple promoter of VSG gene transcription, restricting transcription to that gene which, through genomic rearrangement, is placed downstream from the array. Leader and structural gene sequences are presumably juxtaposed by RNA splicing. Here we show that several trypanosomatids, both those which undergo antigenic variation (*Trypanosoma congolense* and *Trypanosoma vivax*) and those which do not (*T. cruzi* and *Leishmania collosoma*), contain reiterated sequences homologous to the *T. brucei* spliced leader (SL). These results suggest that the SL, although utilized in VSG gene expression, is an ancestral sequence also used in the expression of other trypanosomatid genes.

We have examined several genera of Trypanosomatidae and several species of *Trypanosoma* for sequences homologous to the *T. brucei* SL by Southern analyses. Figure 1 shows restriction enzyme-digested DNAs hybridized with an end-labelled probe complementary to 22 residues of the 35 nucleotide SL. The probe hybridizes to 1.4 kilobase (kb) fragments in each of the *T. brucei* subspecies (lanes 2-4) and to a 1.5-kb fragment in the closely related species *T. equinum* (lane 1). Most of these 1.4 (or 1.5) kb units, each containing one SL sequence, are directly and tandemly repeated^{9,10}. In lanes 5-7 respectively the probe detects one major fragment in DNAs from *T. vivax* (770 base pairs, bp), *T. congolense* (840 bp) and *L. collosoma* (640 bp), while in the complete digest of *T. cruzi* DNA (lane 8) it hybridizes to a series of fragments whose sizes are multiples

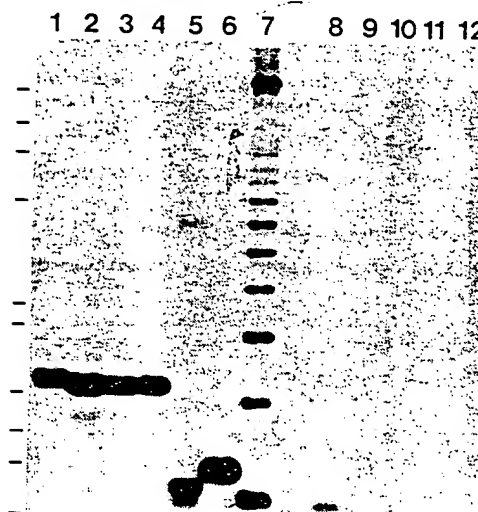


Fig. 1 Identification of genomic sequences homologous to the *T. brucei* VSG mRNA SL in several genera of Trypanosomatidae and several species of *Trypanosoma*. Lane 1, *T. equinum*, IHRI 25; lane 2, *T. b. gambiense*, Treu 1257, VAT L217; lane 3, *T. b. rhodesiense*, ILRAD 1799; lane 4, *T. b. brucei*, IsTaR 1, VAT 1.1118; lane 5, *T. vivax*, ILRAD 1819, SstII; lane 6, *T. congolense*, ILRAD 1180, PvuII; lane 7, *T. cruzi*; lane 8, *L. collosoma*, ATCC 30261; lane 9, *H. muscarum*, ATCC 30260; lane 10, *P. davidi*, ATCC 30287; lane 11, *L. tropica*, NIH strain 173; lane 12, *C. fasciculata*, ATCC 11745. The hash marks indicate λ -HindIII and Φ X174-HaeIII DNA size standards of 23, 9.6, 6.6, 4.4, 2.3, 2.0, 1.4, 1.1, 0.87 and 0.60 kb respectively.

Methods: DNA was isolated¹⁶, cleaved with the indicated restriction enzyme, size fractionated on a 1.0% agarose gel and transferred to a nitrocellulose membrane. This blot was hybridized with a synthetic 5'-[³²P]-end-labelled probe⁹ complementary to 22 of the 35-nucleotide *T. brucei* SL. Hybridization was for 18 h at 37 °C in 5×SSPE (0.9 M NaCl, 0.05 M NaH₂PO₄, 0.005 M Na₂ EDTA, pH 7.4), 0.25% Sarkosyl with 2×10⁶ c.p.m. of probe per ml of hybridization solution. After hybridization, the blot was washed twice at 20 °C and twice at 37 °C in the same solution. All restriction digests were with *TaqI* except where indicated.

of 650 bp. Longer autoradiography of lanes 9-12 (not shown) reveals slight but discrete hybridization to *Herpetomonas muscarum*, *Phytomonas davidi* and *Leishmania tropica* DNAs, indicating that these genomes also contain a sequence related to the *T. brucei* SL.

Surprisingly, two trypanosomatids which do not undergo antigenic variation and which are only distantly related to *T. brucei*, namely *T. cruzi* (Fig. 2a) and *L. collosoma* (Fig. 2b), show the most related organization of SL sequences. Lane 1 of Fig. 2a, b shows complete DNA digests with *BanII* and *TaqI* respectively. The SL probe hybridizes to a 650-bp fragment in *T. cruzi* and to a 640-bp fragment in *L. collosoma*. Partial digestion with these enzymes (Fig. 2a, b, lane 2) reveals a ladder of fragments whose sizes are multiples of the monomer size, demonstrating that, as in *T. brucei*, the monomers are directly and tandemly repeated. When these DNAs are digested with enzymes which do not cleave the monomer (null restriction, lane 3) hybridization is primarily to the intact tandem array, although several smaller fragments are also detected. In *T. brucei*⁹ the majority of such fragments are orphans¹¹, that is, dispersed from the array of tandem repeats.

The organization of SL sequences is more complex in *T. vivax* and *T. congolense*. Complete *SstII* digestion of *T. vivax* DNA (Fig. 2c, lane 1) produces a 770-bp fragment plus several fragments which hybridize less strongly with the SL probe. At least some of these 770-bp fragments are organized in direct tandem repeats since partial *SstII* digestion (Fig. 2c, lane 2) liberates multiples of the 770-bp unit up to the heptamer, at which point the pattern is obscured by other fragments. In addition to larger fragments, the complete *HinfI* digest (Fig. 2c, lane 3) liberates

THIS PAGE BLANK (USPTO)

REFERENCES

- 1 R. Axen, J. Porath and S. Ernback, *Nature*, 214 (1967) 1302.
- 2 M. Wilchek, T. Oka and Y. Topper, *Proc. Natl. Acad. Sci., U.S.A.*, 71 (1974) 1630.
- 3 J. Kohn and M. Wilchek, in W.H. Scouten (Editor), *Solid Phase Biochemistry*, Wiley, New York, 1983, p. 599.
- 4 T.C.J. Gribnau, Ph.D. Thesis, University of Nijmegen, Nijmegen, 1977.
- 5 K. Nilsson and K. Mosbach, *Biochem. Biophys. Res. Commun.*, 102 (1981) 449.
- 6 K. Nilsson and K. Mosbach, *Acta Chem. Scand.*, 35 (1981) 19.
- 7 K. Mosbach, *Pharmacia*, patent pending.
- 8 W.H. Scouten and W. van den Tweel, in I. Chaiken, M. Wilchek and I. Parikh (Editors), *Affinity Chromatography and Biological Recognition*, Academic Press, New York, 1983, p. 299.
- 9 W.H. Scouten and W. van den Tweel, *Ann. N.Y. Acad. Sci.*, 434 (1985) 249.
- 10 W.H. Scouten, M. Dekker and H. Kranenberg, in K. Mosbach (Editor), *Methods Enzymol.*, in press.
- 11 W.H. Scouten, patent application pending.
- 12 P. Sartori and G. Bauer, *J. Fluorine Chem.*, 12 (1978) 203.
- 13 I.L. Knunyants and G.A. Sokol'skii, *Angew. Chem. Int. Ed. Engl.*, 11 (1972) 7.
- 14 D.C. England, M.A. Dietrich and R.V. Lindsey, Jr., *J. Am. Chem. Soc.*, 82 (1960) 6181.
- 15 D.C. England and H. Oak, U.S. Pat., 2,852,554 (1958).

Journal of Chromatography, 376 (1986) 299-305
Biomedical Applications
 Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMBIO. 2977

DIVINYLSULPHONE-ACTIVATED AGAROSE

FORMATION OF STABLE AND NON-LEAKING AFFINITY MATRICES BY IMMOBILIZATION OF IMMUNOGLOBULINS AND OTHER PROTEINS

A. LIHME, C. SCHAFER-NIELSEN, K.P. LARSEN, K.G. MÜLLER and
 T.C. BØG-HANSEN*

The Protein Laboratory, University of Copenhagen, Sigurdsgade 34, DK-2200 Copenhagen N (Denmark)

SUMMARY

Divinylsulphone-activated agarose is an attractive alternative to several of the activated supports usually used. Unlike CNBr-activated gels, it does not leak the immobilized protein at high pH. It reacts readily with proteins at near-neutral pH (unlike the epoxy-activated supports). Generally, divinylsulphone-activated agarose reacts with amino, hydroxyl, and sulphhydryl groups, thus allowing immobilization of a wide spectrum of ligands. Moreover, it is available in an aqueous suspension free of organic solvents and neither requires time-consuming swelling nor washing.

INTRODUCTION

Cross-linking and activation of agarose with divinylsulphone (DVS) was described by Porath and co-workers in 1975 [1, 2] but has achieved little recognition since then, even though the method has been applied to the immobilization of carbohydrates [3] and hormones [4].

Our preliminary investigations lead us to believe that this method for the immobilization of ligands for affinity chromatography has more useful applications than has been recognized so far. Thus, it has been our aim to investigate the coupling parameters of DVS-derivatized agarose gel as well as the functional integrity of coupled ligands, with special reference to immunosorption techniques. We conclude that DVS-activated agarose is a useful matrix for the immobilization of biomolecules, comparing favourably with CNBr- and epoxy-activated gels.

THIS PAGE BLANK (USPTO)

EXPERIMENTAL

Materials

DVS-activated agarose (MINI-LEAK), horseradish peroxidase (HRP), crude goat antiserum against rabbit immunoglobulin G (IgG) and lectins concanavalin A (Con A), wheatgerm agglutinin (WGA), *Lens culinaris* agglutinin (LCA), *Pisum sativum* agglutinin (PSA), soya bean agglutinin (SBA) peanut agglutinin (PNA), *Solanum tuberosum* agglutinin (STA) and *Vicia villosa* agglutinin (VVA) were a gift from Kem-En-Tec Biotechnology Corp. (Hellerup, Denmark). Bovine albumin and bovine γ -globulins were from Sigma (St. Louis, MO, U.S.A.). CNBr-activated agarose was from Pharmacia (Uppsala, Sweden) as was CNBr-activated Sepharose 6B. Rabbit and human immunoglobulins were prepared on immobilized protein A according to standard procedures. Human IgG was labelled with ^{125}I by the chloramine-T method.

Methods

For experimental conditions, see legends to Figs. 1-4 and Table I.

RESULTS

Coupling as a function of ionic strength

Coupling of proteins was found to proceed in near-neutral to weakly alkaline media from pH 7.5 to 9.5, coupling faster the more alkaline the pH. In the following experiments, as well as for routine coupling, we chose to couple at pH 8.6. All coupling procedures were performed at room temperature for 14-20 h.

We found that various proteins couple with different efficiencies depending upon the ionic strength of the coupling medium. Thus, as indicated in Fig. 1,

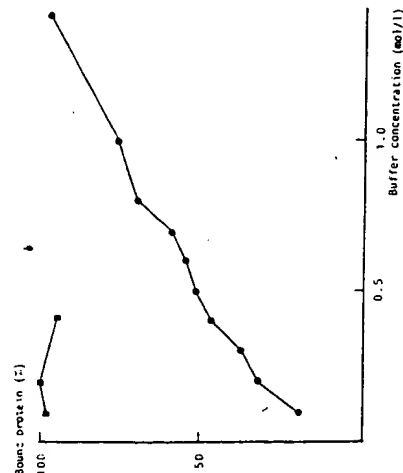


Fig. 1. Coupling of pure proteins to DVS-activated agarose. (●) Bovine γ -globulins; (●) bovine serum albumin. The coupling was performed at pH 8.6 for 14 h at room temperature at a concentration of 15 mg/ml protein, corresponding to 30 mg of protein per g of wet gel. The coupling is expressed as the percentage of bound protein as a function of the concentration of potassium phosphate buffer, pH 8.6 (0.05-1.4 M).

albumin only coupled efficiently to the gel at high ionic strength, while the coupling of γ -globulins showed no dependence on ionic strength.

Coupling capacity and yield

As the coupling capacity and coupling yield are important parameters in practical work with affinity chromatographic procedures, we investigated these parameters as a function of the amount of added protein. We defined the coupling capacity as the maximal amount of protein that can be coupled to 1 ml of wet, sedimented gel, and the coupling yield as the percentage of the total amount of added protein that is coupled. Bovine albumin and bovine γ -globulin were used as model proteins, and for comparison we performed parallel experiments with CNBr-activated Sepharose (Fig. 2). We used optimal conditions for the coupling, e.g. high ionic strength for albumin and moderate ionic strength for γ -globulins when coupling to DVS-activated gel. Moderate ionic strength was used when coupling to CNBr-activated gel, as coupling of albumin to this gel is independent of ionic strength (over a broad range). As seen in Fig. 2, the DVS-activated gel has a higher capacity for albumin at high albumin concentration than the CNBr-activated gel, whereas the two gels show similar capacity for binding of γ -globulins. We were pleased to find that the DVS-activated gel couples with yields better than 80%, up to a load as high as 80 mg of protein per ml of activated gel.

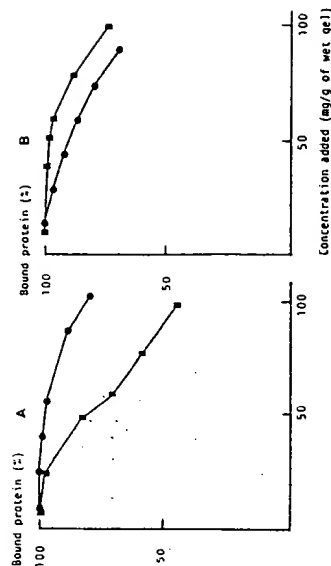


Fig. 2. Coupling of pure proteins to DVS-activated agarose (●) and to CNBr-activated Sepharose 6B (○). The yield of bound protein (%) is given as a function of the amount of protein added per g of wet activated gel. (A) Coupling of bovine serum albumin: the coupling buffer was 1.4 M potassium phosphate (pH 8.6) for DVS-activated agarose and 0.2 M potassium phosphate (pH 8.6) for CNBr-activated Sepharose 6B. (B) Coupling yield of bovine γ -globulin: the coupling buffer was 0.2 M potassium phosphate (pH 8.6). During coupling, the protein concentration was 20-30 mg/ml of coupling suspension. The coupling was performed for 14 h at room temperature.

Stability of the coupled gel

As high and low pH values are often applied for elution of immunosorption columns and, generally, with deforming buffers in affinity chromatography, we examined the leakage of immobilized protein at pH 2.5-11.5. For this experiment two different proteins, HRP and radioactively labelled human IgG, were immobilized on DVS- and CNBr-activated gels.

THIS PAGE BLANK (USPTO)

Fig. 3. Preparation of immunosorbed antibodies. (A) Crossed immunoelectrophoresis of 2 μ l of rabbit serum with 300 μ l of crude goat antiserum against rabbit IgG (arrow indicates IgG). (B) Crossed immunoelectrophoresis of 2 μ l of rabbit serum with 300 μ l of eluted antibody fraction. After elution at pH 11.5, the antibody fraction is formed by anti-IgG. First-dimension anode is to the right (electrophoresis for 1 h at 10 V/cm); second-dimension anode on top (electrophoresis for 16 h at 1.8 V/cm). Buffer was 1% Tris-veronal (pH 8.6) Lite agarose HSA. Plate dimensions 7 x 10 cm.



TABLE I

PROTEIN LEAKING

HRP: purified HRP was coupled to DVS- and CNBr-activated gels at a concentration of 5 mg/g of wet sedimented gel, washed with high- and low-pH buffers and then incubated for 15 h at pH 11.5 and 2.5. Release of HRP from the gels was determined continuously by spectrophotometry at 403 nm and calculated as HRP liberated (ng/min) per mg HRP immobilized. IgG: purified, radioactively labelled human IgG was immobilized on the two gels at a concentration of 10 mg/g wet sedimented gel, washed with high- and low-pH buffers and then incubated for 24 h at pH 11.5, 7.0 and 2.5. Release of radioactive protein was determined and calculated as IgG liberated (ng/min) per mg IgG immobilized.

Protein released	HRP (ng/min mg)			IgG (ng/min mg)		
	pH 11.5	pH 2.5	pH 11.5	pH 7.0	pH 2.5	
From DVS-activated gel	0.2	< 0.1	1.5	0.35	1.1	
From CNBr-activated gel	11.0	2.8	36.8	0.35	1.4	

The leakage from the two gels was calculated from the absorbance at 403 nm owing to released HRP and from radioactivity of the released immunoglobulin. Table I shows that the DVS-activated gel gives coupled products that are at least one order of magnitude more stable at alkaline pH than the products of the CNBr-activated gel.

Applications in affinity chromatography

The high coupling capacity, combined with the high stability of the chemical bonds, makes DVS-activated gel a potentially useful matrix in affinity chromatography. In particular, the stability as expressed by the comparatively low leakage at high pH values makes this gel an attractive alternative for immunosorbent techniques.

Coupling of antigen for immunosorption: immobilized rabbit immunoglobulin for preparation of goat anti-rabbit immunoglobulin. For the preparation of affinity-purified goat anti-rabbit immunoglobulin from crude oligospecific antiserum, we immobilized pure rabbit immunoglobulins to DVS-activated agarose to a final concentration of 15 mg of rabbit immunoglobulins per g of wet gel. The matrix obtained was used in an immunosorbent column, which was loaded with the crude antiserum, washed with saline and eluted with 0.1 M phosphate buffer (pH 11.5). Fig. 3A illustrates the oligospecificity of the crude antiserum when used in the second dimension of a crossed immunoelectrophoresis of rabbit serum. The crude antiserum precipitates a number of rabbit proteins in addition to IgG. In Fig. 3B the non-binding fraction, the run-through, is used as antibody in the second dimension. As seen from the disappearance of the IgG precipitate, all anti-IgG has been adsorbed to the column. After desorption at pH 11.5, the desorbed antibody fraction, the non-sense goat anti-rabbit IgG, showed only activity against IgG (Fig. 3C).

When loading the column at its maximal capacity, 10 ml of crude antiserum could be completely adsorbed for anti-IgG per ml of gel. On a 2-ml column we could prepare 28 mg of affinity-purified goat anti-rabbit IgG. These data indicate that the immobilized rabbit IgG, to a large extent, is sterically available and active as antigen on the matrix.

THIS PAGE BLANK (USPTO)

Coupling of antibodies for immunosorption: goat anti-rabbit immunoglobulin immobilized for preparation of rabbit immunoglobulin. To investigate the performance of antibodies coupled to DVS-activated agarose, we immobilized the non-sense affinity-purified goat anti-rabbit IgG (described above) to a final density of 1 mg of antibody per ml of gel.

A column with 10 ml of this antibody gel was prepared and, after careful washing with high- and low-pH buffers (pH 2.2 and 11.5, respectively), the column was loaded with rabbit serum. After washing with saline, the column was eluted with 0.1 M glycine hydrochloride (pH 2.2) and the eluate was immediately neutralized. When loaded to its maximal capacity, the column could yield between 4 and 5 mg of pure rabbit IgG repeatedly, without loss of activity. Fig. 4 shows an analysis of the rabbit serum and the eluted immunoglobulin by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS PAGE). Only IgG was detected in the eluate; other rabbit proteins could not be seen (column 1).

Our data indicate that between 40 and 50% of the DVS-activated agarose-immobilized antibody is exposed on the matrix in a sterically available and active form.

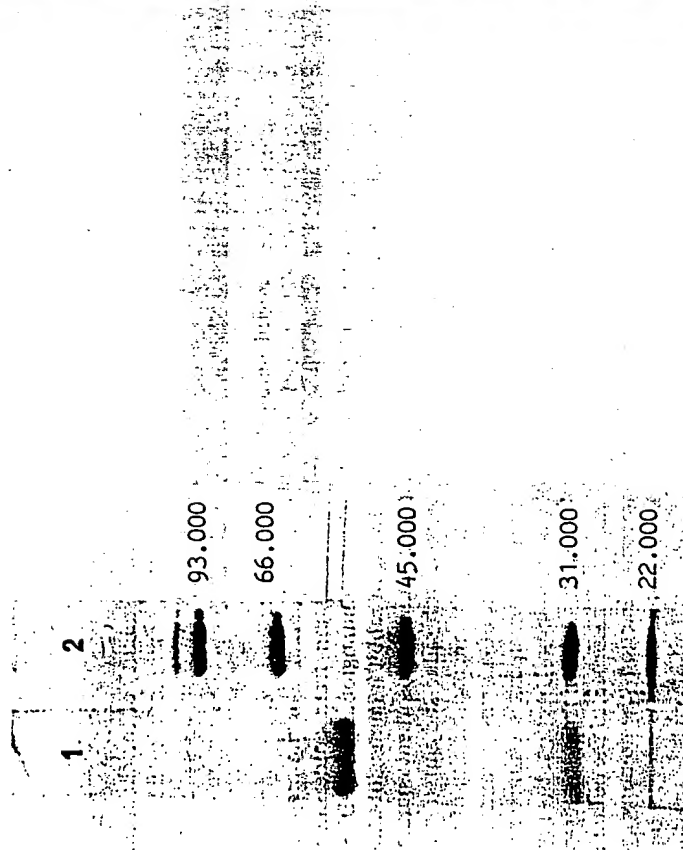


Fig. 4. Preparation of rabbit immunoglobulin. Analysis by SDS-PAGE. Column 1: the eluted immunoglobulin fraction with heavy and light chains; electrophoresis was performed at 30 mA overnight. Column 2: a set of marker proteins with the molecular weights given (Kem-En-Tec).

DISCUSSION

The DVS-activated agarose appears to exhibit an extraordinarily high reactivity when coupling immunoglobulins, as compared to albumin (Fig. 1). In other experiments, we have been able to demonstrate that this selectivity is also expressed during immobilization of complex protein mixtures as crude antiserum, mouse ascites fluid and tissue culture supernatant (in preparation). The selective adsorption of immunoglobulins to the mercaptoethanol derivative of DVS-activated agarose (the so-called T-gel [7]) was presented at this meeting by Porath. This adsorption is dependent upon the ionic strength and may be based upon a structural interaction between a surface area of the immunoglobulin molecule and the C-S-C configuration. This promising adsorbent can be prepared easily by incubating DVS agarose with mercaptoethanol at slightly alkaline pH [7].

To assess the activity of another group of binding proteins after coupling to DVS-activated agarose, we chose to immobilize lectins. This was done at a concentration of 5–10 mg of lectin per ml of gel and the immobilized lectins (see *Materials*) were found to be fully active and able to separate human serum proteins into a binding fraction and a non-binding fraction (in preparation).

In contrast to CNBr-activated agarose, the DVS-activated agarose also reacts with hydroxyl groups. We immobilized L-fucose on DVS-activated agarose at pH 11 for 24 h at room temperature [5]. We then used this gel successfully as an affinity adsorbent for purification of the lectin UEA-I from crude extracts of *Ulex europaeus* seeds. The capacity of this gel was determined to be 4–5 mg of lectin per ml of gel [5], which is more than three times as much as reported earlier [6].

ACKNOWLEDGEMENT

The Harboe Foundation is thanked for generous support of the Protein Laboratory.

REFERENCES

- 1 J. Porath, T. Lfås and J.-Ch. Janson, *J. Chromatogr.*, **103** (1975) 49–62.
- 2 L. Sundberg and J. Porath, *Protides Biol. Fluids, Proc. Colloq.*, **23** (1975) 517–523.
- 3 H.J. Allen and E.A.Z. Johnson, *Carbohydr. Res.*, **58** (1975) 253–265.
- 4 M.R. Saitam, *J. Chromatogr.*, **215** (1981) 143–152.
- 5 K.G. Muller, C. Schafer-Nielsen and A. Lihme, in T.C. Bøg-Hansen and E. van Driessche (Editors), *Lectins, Biology, Biochemistry, Clinical Biochemistry*, Vol. 5, Walter de Gruyter, Berlin, 1986, pp. 135–138.
- 6 H.J. Allen and E.A.Z. Johnson, *Carbohydr. Res.*, **58** (1977) 253–265.
- 7 J. Porath, F. Maisano and M. Belew, *FEBS Lett.*, **185** (1985) 306–311.

THIS PAGE BLANK (USPTO)

Involvement of the Golgi apparatus in the secretion of α -amylase from gibberellin-treated barley aleurone cells

F. Gubler¹*, J.V. Jacobsen² and A.E. Ashford¹

¹ School of Botany, University of New South Wales, P.O. Box 1, Kensington, N.S.W. 2033, and

² CSIRO Division of Plant Industry, Canberra, A.C.T. 2601, Australia

Abstract. Localisation of α -amylase (EC 3.2.1.1) in barley aleurone cells treated with gibberellic acid has been achieved using protein A-gold-labelled polyclonal antibodies. Gold particles were located almost exclusively over the lumen of the rough endoplasmic reticulum and cisternae of the Golgi apparatus. The label was most concentrated over the Golgi apparatus. This indicates that the Golgi is involved in the secretion of α -amylase protein from aleurone cells.

Key words: Aleurone - α -Amylase - Gibberellin and amylase secretion - Golgi apparatus - *Hordeum* (α -amylase secretion) - Protein secretion.

Introduction

In all cases examined, the intracellular transport pathway for protein destined to be secreted from animal cells involves the Golgi apparatus (see Farquhar and Palade 1981 and Kelly 1985 for reviews). In higher-plant cells, although the Golgi is known to be involved in transport of newly synthesised storage protein into protein bodies in bean cotyledons (see Greenwood and Chrispeels 1985), there is no direct evidence of Golgi involvement in transport of protein that is secreted from the cell. This lack of evidence arises primarily from the scarcity of plant cells which secrete large amounts of protein and are well characterised.

The best-studied protein-secreting tissue of plants is the barley aleurone layer, which is stimulated by gibberellic acid (GA_3) to synthesise and secrete a number of hydrolytic enzymes, that are normally involved in mobilisation of endosperm

reserves during germination (for recent reviews, see e.g. Ashford and Gubler 1984; Jacobsen 1984). The synthesis of α -amylase, the most abundant of these proteins which comprises some 50% of newly synthesised protein, has been intensively studied and there is now evidence that GA_3 regulates expression of the α -amylase genes at the level of transcription (Jacobsen and Beach 1985). However, the intracellular transport pathway for α -amylase is still controversial. The unprocessed α -amylase polypeptide contains a signal sequence (Rogers and Milliman 1983; Chandler et al. 1984) and pulse-chase experiments have located newly synthesised protein in the microsomal fraction with a latency that indicates that the enzyme may be sequestered in the lumen of the endoplasmic reticulum (ER) (Jones and Jacobsen 1982). Both observations taken together indicate that the ER is involved in the α -amylase transport pathway, but the route taken by the enzyme beyond this point is unclear. At least four possible pathways have been suggested, involving transfer via the cytosol (Jones 1972), lysosomes (Gibson and Paleg 1976), ER-derived vesicles (Vigil and Ruddat 1973; Locy and Kende 1978; Pylotis et al. 1979) or Golgi bodies (Jones 1969; Fernandez and Staehelin 1985). However, none of these proposals has been supported by direct evidence.

We have localised α -amylase in situ in sections of GA_3 -treated aleurone embedded at low temperature in Lowicryl K4M (Carlemalm et al. 1982), using protein A-gold-labelled anti- α -amylase antibodies. This has enabled us to determine in which organelles the enzyme is located.

Material and methods

Preparation and incubation of aleurone layers. Aleurone layers were prepared from grains of *Hordeum vulgare* L. cv. Himalaya

* To whom correspondence should be addressed

Abbreviations: ER = endoplasmic reticulum; GA_3 = gibberellic acid; PBS = phosphate-buffered saline



THIS PAGE BLANK (USPTO)

(Washington State University, Pullman, Wash., USA: 1979 harvest) and incubated with 10^{-6} M GA_3 in 0.01 M $CaCl_2$ at 25° C as described in Chrispeels and Varner (1967) except that the acetate buffer was omitted (see Jacobsen and Higgins 1982).

Preparation and purification of antibodies. Barley α -amylase was purified from the incubation medium of GA_3 -treated aleurone layers as described by Jacobsen and Higgins (1982). Rabbit antibodies were prepared against the α -amylase and purified by affinity chromatography on an α -amylase-Sepharose 4B column as previously described (Higgins et al. 1982; Jones and Jacobsen 1982). Immunoreactivity towards the low- and high-isoelectric-point groups of α -amylase isoenzymes was tested by the double-immunodiffusion technique (Jacobsen and Higgins 1982).

Western blot. The specificity of the antibodies was checked by the Western-blot technique of Towbin et al. (1979). An extract was prepared from isolated, 20-h GA_3 -treated aleurone layers, by homogenising the layers in 10^{-2} M sodium-acetate buffer (pH 4.8) containing 10^{-2} M $CaCl_2$ (Chrispeels and Varner 1967) and the insoluble fraction was removed by centrifugation at 10000 g for 5 min. This extract and purified α -amylase were each fractionated by gradient sodium dodecyl sulphate-polyacrylamide gel electrophoresis (12.5–25% acrylamide), transferred to nitrocellulose, and stained with anti- α -amylase antibodies ($5 \mu\text{g} \cdot \text{ml}^{-1}$) followed by protein A-peroxidase (Towbin et al. 1979; Bittner et al. 1980). Staining specificity was checked by a control, substituting the α -amylase antibodies with rabbit anti-sheep immunoglobulin G serum, 1-in-200 dilution (supplied by D. Spencer, CSIRO, Division of Plant Industry).

Fixation and low-temperature embedding. Small strips ($1 \times 2 \text{ mm}^2$) of isolated barley aleurone layers treated with GA_3 for 16 h were fixed in 2% glutaraldehyde in 0.03 M 1,4-piperazinediethanesulfonic acid (Pipes) buffer (pH 7.5) and 0.01 M $CaCl_2$ for 3 h at 0° C. After fixation the tissue was dehydrated in an ethanol series (25, 50, 75, 95, 100%) with progressive lowering of temperature to a final temperature of -20°C at 100%, and then infiltrated with Lowicryl K4M (Chemische Werke Lowi, Waldkraiburg, FRG) for 9 d at -20°C . Ratios of Lowicryl K4M components were 4 g crosslinker A:26 g monomer B:0.15 g initiator C. Tissue was embedded in Lowicryl K4M in containers immersed in a liquid refrigerant bath, at a resin temperature of approx. -45°C . Polymerisation was brought about by ultraviolet light in an atmosphere of nitrogen and special care was taken to prevent the large temperature rise which occurs when Lowicryl K4M is polymerised under inadequate cooling conditions (Ashford et al. 1986).

Immunocytochemistry. The protein A-gold technique was used for the immunocytochemical localisation of α -amylase on thin sections (silver-gold). Monodisperse colloidal gold (15 nm) was prepared by reduction of chloroauric acid (BDH Chemicals, Poole, UK) with trisodium citrate (Frens 1973) and was complexed with protein A (Pharmacia, Uppsala, Sweden) by the method of Roth et al. (1978). Sections were mounted on carbon-formvar-coated copper grids and were pretreated by floating on drops of phosphate-buffered saline (PBS: 0.01 M sodium-phosphate buffer, pH 7.2, and 0.15 M NaCl with $0.2 \text{ mg} \cdot \text{ml}^{-1}$ NaN_3) containing $20 \text{ mg} \cdot \text{ml}^{-1}$ ovalbumin for 10 min. They were then incubated (again by floating on drops) in PBS-Tween (0.01 M sodium-phosphate buffer, pH 7.2, 0.5 M NaCl and 0.2% polyoxyethylenesorbitan monolaurate [Tween 20] with $0.2 \text{ mg} \cdot \text{ml}^{-1}$ NaN_3) containing $20 \mu\text{g} \cdot \text{ml}^{-1}$ α -amylase antibody and $20 \text{ mg} \cdot \text{ml}^{-1}$ ovalbumin for 1 h. Following incuba-

tion, the sections were rinsed by repeated immersion in PBS-Tween for 5 min, transferred on to a drop of $20 \text{ mg} \cdot \text{ml}^{-1}$ ovalbumin in PBS for 10 min, and then incubated on a drop of solution containing protein A-gold diluted in $5 \text{ mg} \cdot \text{ml}^{-1}$ ovalbumin in PBS ($A_{525}=0.05$) for 1 h. Sections were then rinsed in PBS followed by distilled water and allowed to dry. All steps were done at 20° C. The sections were post-stained with 2% aqueous uranyl acetate followed by 1.3% lead citrate, and examined in a Philips 300 electron microscope (Philips, Eindhoven, Netherlands) at 80 kV.

Specificity of staining was evaluated by the following controls. Sections were incubated with: 1) α -amylase antibody which had been previously absorbed with 50-fold excess of antigen, followed by protein A-gold; 2) rabbit anti-goat immunoglobulin G antibodies (Kirkegaard & Perry Laboratories, Gaithersburg, Md., USA), followed by protein A-gold; 3) buffer (without antibody), followed by protein A-gold.

Quantitation of immunolabelling. The labelling density over various cellular organelles was measured by counting the number of gold particles per unit area of that organelle profile in immunolabelled sections. The area of individual organelle profiles was measured with an Apple graphics tablet (Apple Computer, Cupertino, Cal., USA) connected to an Apple IIe computer and a carbon grating replica ($2160 \text{ lines} \cdot \text{mm}^{-1}$; Agar Aids, Stansted, UK) was recorded for calibration. Sections were taken from 2 aleurone layers (aleurone 1 and 2) and the data were recorded from a total of 135 micrographs each at 15600 \times magnification.

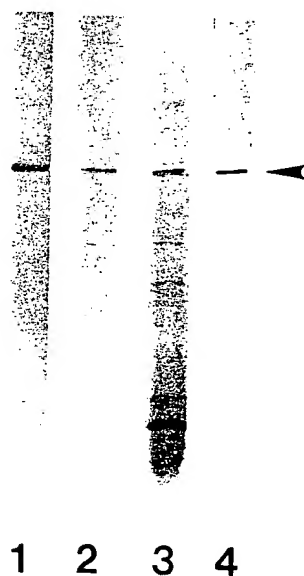


Fig. 1. Analysis of antibody specificity by the Western-blot technique. Purified barley α -amylase and extracts from GA_3 -treated barley aleurone layers were immunoblotted with affinity-purified anti- α -amylase antibodies, after sodium dodecyl sulphate-polyacrylamide gel electrophoresis and transfer to nitrocellulose. Purified α -amylase: lane 1, stained with amido black; lane 2, labelled with anti- α -amylase antibodies. Aqueous extract of GA_3 -treated aleurone layers: lane 3, stained with amido black; lane 4, labelled with anti- α -amylase antibodies. Only a single band (arrow) was detected in the aleurone extract when blotted with anti- α -amylase antibody. This band co-migrated with α -amylase. In a control where replicates of lanes 2 and 4 were reacted with labelled rabbit anti-sheep immunoglobulin-G serum there was no staining (not shown).

THIS PAGE BLANK (USPTO)

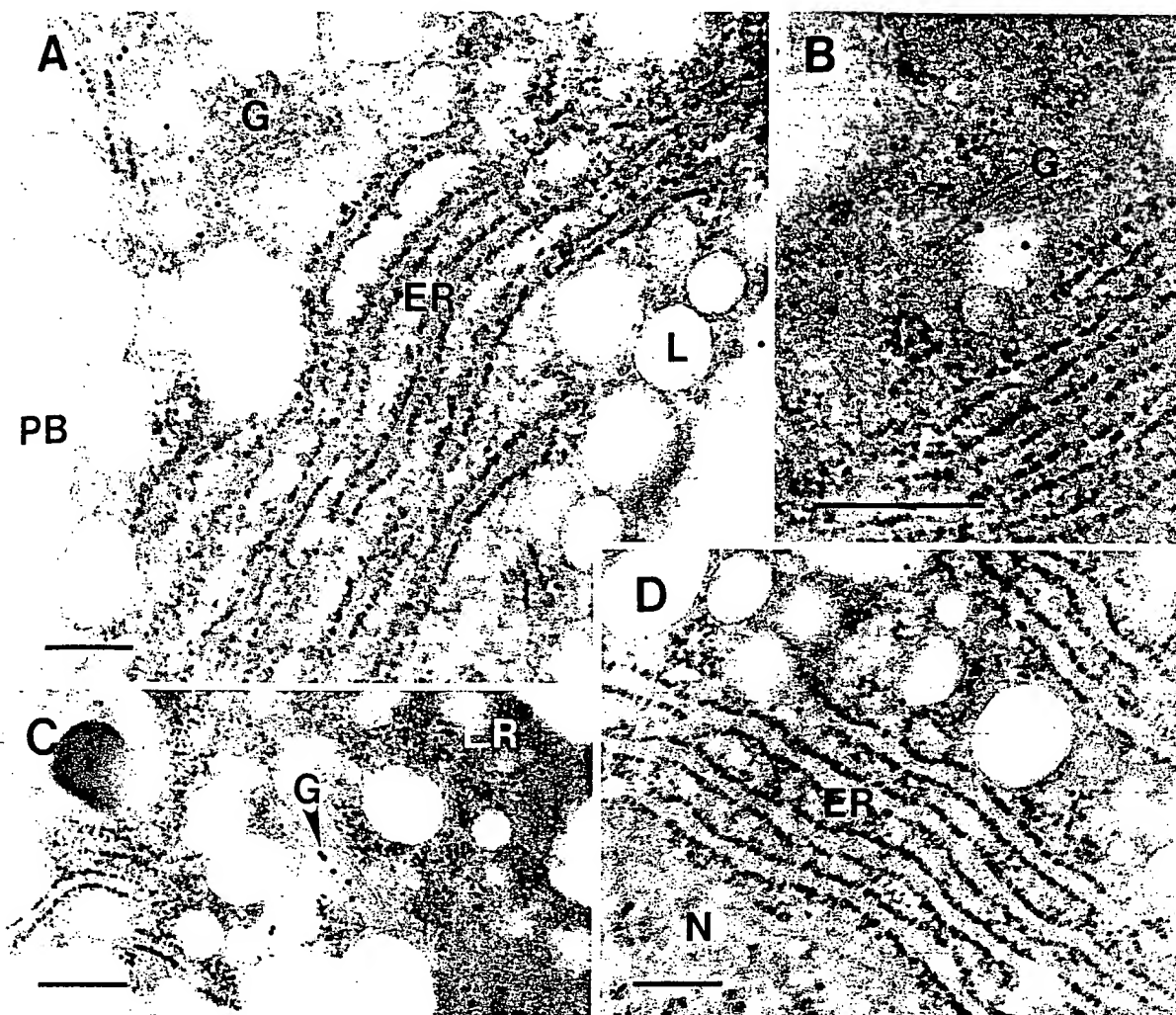


Fig. 2A-D. Immunocytochemical localisation of α -amylase in barley aleurone cells treated with GA_3 for 16 h. Lowicryl K4M-embedded sections were stained for α -amylase using the protein A-gold technique. Because of the absence of post-fixation with osmium tetroxide, the membranes appear negatively stained. A The GA_3 -treated aleurone cells contained extensive rough ER (ER). Golgi bodies (G) were also common especially around the cell periphery. In this micrograph labelling is specifically associated with the rough ER lumen and regions adjacent to the Golgi body. Other organelles such as protein bodies (PB) and lipid bodies (L) had little or no label. B and C Micrographs of Golgi bodies (G) showing labelling over the cisternae. Note label over region of rough ER sectioned tangentially (ER). D Control section stained with only protein A-gold showing no labelling over the rough ER (ER). Note the non-specific staining of the nucleus (N). A, D $\times 44000$; B $\times 78000$; C $\times 46000$. Bars = $0.3 \mu m$

Results

The specificity of the affinity-purified antibodies produced against barley aleurone α -amylase was tested by the Western-blot technique (Fig. 1). When blotted against a crude aleurone extract from GA_3 -treated cells, the antibodies reacted only with the α -amylase band. They were shown by double immunodiffusion tests to be immunoreactive against both low- and high-isoelectric-point groups of α -amylase isoenzymes (data not illustrated), as had been found by Jacobsen and Higgins (1982).

The preservation of GA_3 -treated aleurone cells, embedded at low temperature in Lowicryl K4M, was generally good. The ultrastructure was characteristic for GA_3 -treated tissue with extensive areas of rough ER and of wall digestion (Jones 1969). The membranes appeared negatively stained but the rough ER was readily recognisable by the rows of ribosomes. The immunocytochemical localisation of α -amylase in GA_3 -treated aleurone cells using protein A-gold is shown in Fig. 2. High levels of labelling were present over the rough ER and Golgi bodies. In contrast, other organelles had very low labelling. In Fig. 2A labelling over the

THIS PAGE BLANK (USPTO)

Table 1. Density of immunolabelling over organelles in GA₃-treated aleurone cells (gold particles per $\mu\text{m}^2 \pm \text{SD}$)

	Aleurone layer 1		Aleurone layer 2	
	Amylase antibody	Antigen-absorbed control ^a	Amylase antibody	Antigen-absorbed control ^a
Golgi	120 \pm 47	<1	100 \pm 37	<1
Rough ER	44 \pm 19	<1	51 \pm 15	1.8 \pm 2.8
Nucleus	4.5 \pm 1.0	—	3.8 \pm 1.5	—
Microbody	3.0 \pm 2.2	<1	1.4 \pm 0.8	<1
Lipid body	<1	—	≤ 1	—
Plastid	≤ 1	—	<1	—
Protein body	<1	—	<1	—
Mitochondrion	<1	—	<1	—

^a α -Amylase + α -amylase antibody; — = not counted

rough ER is shown to be both in the lumen and associated with the membranes. Gold particles were rarely seen in the cytoplasmic areas between ER cisternae. The labelling of Golgi involved all cisternae as well as regions adjacent to the *cis* and *trans* faces (Fig. 2A–C). Occasionally vesicles with high levels of labelling were observed but none in close proximity to Golgi bodies. Control treatments testing the specificity of the reactions all had very low labelling (e.g. Fig. 2D).

Although the distribution of label is clear from the micrographs shown, in any labelling technique, proper evaluation of differences in levels between organelles always requires replicate sampling and quantitative data (Stoward 1985). Quantitative estimates of the labelling densities over various cellular compartments confirmed the concentration of label in the rough ER and Golgi bodies (Table 1). Density of labelling over most organelles was about one gold particle per $1 \mu\text{m}^2$ with the labelling in the microbodies and nuclei somewhat higher. Labelling associated with the rough ER was at least ninefold higher, and in the Golgi 25-fold higher than that over any other cellular site measured. There was thus two to three times more label over the Golgi than over the rough ER. In contrast, the labelling in these organelles in the antigen-absorbed control was at the level of the background.

Discussion

Localisation of α -amylase by protein A-gold immunocytochemistry almost exclusively in the rough ER and Golgi apparatus provides direct evidence for the involvement of these organelles in the intracellular pathway of α -amylase secretion.

In the interpretation of these data it is important to remember that α -amylase is synthesised *de novo* in response to GA₃ and that all enzyme is destined for secretion (Yomo and Varner 1971). This means that all α -amylase localised in the cell is likely to be at some point along its secretory pathway, in contrast to other enzymes such as acid phosphatase (Pyliotis et al. 1979) where this cannot be determined because this enzyme is normally present in the cells and only a fraction is secreted (Ashford and Gubler 1984).

Localisation of α -amylase within the lumen of the rough ER is consistent with biochemical evidence that isoenzymes from both α -amylase groups carry a leader sequence (Rogers and Milliman 1983; Chandler et al. 1984) and that the lower-molecular-weight form is sequestered in microsome (Jones and Jacobsen 1982). It is also consistent with the proliferation of rough ER observed in aleurone cells treated with GA₃ (Jones 1969; Vigil and Ruddat 1973). All this evidence taken together indicates that synthesis and the initial part of the secretory pathway follow closely that found for animal tissues (see Kreil 1981 for a review) and that the tenets of the signal hypothesis apply here as well.

In animal cells the Golgi complex has long been recognised as playing a central role in the sorting, modification and packaging of newly synthesised protein transported from the rough ER both to extracellular and intracellular destinations (see Farquhar and Palade 1981; Kelly 1985). Evidence is now accumulating for a similar situation in plant cells. It is known from biochemical, ultrastructural and immunocytochemical work that the Golgi apparatus is involved in the processing and packaging of storage proteins and lectins destined for vacuoles or protein bodies (Craig and Goodchild 1984; Herman and Shannon 1984a, b; Vitale and Chrispeels 1984; zur Nieden et al. 1984; Greenwood and Chrispeels 1985). Involvement of the Golgi apparatus in extracellular secretion of plant glycoproteins has been proposed by Gardiner and Chrispeels (1975) and Kawasaki and Sato (1979) on the basis of cell-fractionation experiments. More recent work (Akazawa and Hara-Nishimura 1985; Mitsui et al. 1985) has shown post-translational modification of N-linked oligosaccharides of rice-scutellum α -amylase and partial inhibition of this by monensin, both of which are indicators of Golgi processing. Our immunocytochemical data directly demonstrate the presence of α -amylase in Golgi bodies in GA₃-treated barley aleurone, known to be secreting this enzyme, and clearly implicate the Golgi in the secretion of α -

THIS PAGE BLANK (USPTO)

amylase. Golgi bodies are commonly seen in electron micrographs of aleurone cells (Jones 1969; Fernandez and Staehelin 1985). As Fernandez and Staehelin (1985) pointed out. Golgi bodies in aleurone are not hypertrophied like those in root-cap cells of maize which are secreting carbohydrate (Dauwalder and Whaley 1982), and this, together with an apparent lack of prominence of Golgi bodies compared with rough ER, may explain the lack of attention previously given to this organelle.

The presence of gold labelling over all parts of the Golgi cisternae in the barley-aleurone cells and quantitative data indicating that there is concentration of label over the Golgi compared with the rough ER, support the view that enzyme is transported through the entire Golgi stack and that the Golgi follows the rough ER in the secretory pathway. The most probable pathway of release from the Golgi is via Golgi-derived vesicles as appears to be the case for another secreted enzyme, peroxidase (Ashford and Gubler 1986). Our data show that α -amylase is not stored prior to secretion in zymogen-like granules as found for α -amylase and other enzymes secreted from pancreatic acinar cells (Palade 1975). In this respect the secretory pathway in aleurone is most like that described

"constitutive" in animal cells, where there is only a short delay between synthesis and secretion and enzyme does not accumulate temporarily in the cells (Kelly 1985). The antibody used here is reactive towards both high- and low-isoelectric-point groups of α -amylase (Jacobsen and Higgins 1982). Although we cannot eliminate the possibility that only one type of enzyme has been detected in these studies, it is more likely that both are localised simultaneously by our immunocytochemistry, and therefore that the Golgi apparatus is involved in the secretion of both α -amylase isoenzyme groups.

In conclusion, our data show directly that Golgi is involved in α -amylase secretion from aleurone cells. These findings indicate that protein secretion in plants follows similar pathways to that in animal cells and therefore that there may be a universal mechanism for eukaryotes.

We gratefully acknowledge Dr. D. Spencer for the Western-blot analyses. We would also like to thank Dr. S. Craig for helpful advice and the Electron Microscope Units of the Universities of New South Wales and Sydney for providing facilities. A.E.A. and F.G. thank the Australian Research Grants Scheme for financial support.

References

- Akazawa, T., Hara-Nishimura, I. (1985) Topographic aspects of biosynthesis, extracellular secretion and intracellular storage of proteins in plant cells. *Annu. Rev. Plant Physiol.* **36**, 441-472
- Ashford, A.E., Allaway, W.G., Gubler, F., Lennon, A., Sleepers, J. (1986) Temperature control in Lowicryl K4M and glycol methacrylate during polymerisation - is there a low temperature embedding method? *J. Microsc.*, in press
- Ashford, A.E., Gubler, F. (1984) Mobilization of polysaccharide reserves from endosperm. In: *Seed physiology*, vol. 2: Germination and reserve mobilization. pp. 117-162, Murray, D.R., ed. Academic Press, Sydney
- Ashford, A.E., Gubler, F. (1986) The pathway of secretion of enzymes from isolated barley aleurone layers. In: *IV Int. Symp. on Pre-harvest Sprouting in Cereals*, in press, Mares, D.J., ed. Westview Press, Boulder, Col., USA
- Bittner, M., Kupferer, P., Morris, C.F. (1980) Electrophoretic transfer of proteins and nucleic acids from slab gels to diazobenzoyloxymethyl cellulose or nitrocellulose sheets. *Anal. Biochem.* **102**, 459-471
- Carlemalm, E., Garavito, R.M., Villiger, W. (1982) Resin development for electron microscopy and an analysis of embedding at low temperature. *J. Microsc.* **126**, 123-143
- Chandler, P.M., Zwar, J.A., Jacobsen, J.V., Higgins, T.J.V., Inglis, A.S. (1984) The effects of gibberellic acid and abscisic acid on α -amylase mRNA levels in barley aleurone layers studies using an α -amylase cDNA clone. *Plant Mol. Biol.* **3**, 407-418
- Chrispeels, M.J., Varner, J.E. (1967) Gibberellic acid-enhanced synthesis and release of α -amylase and ribonuclease by isolated barley aleurone layers. *Plant Physiol.* **42**, 398-406
- Craig, S., Goodchild, D.J. (1984) Periodate-acid treatment of sections permits on-grid immunogold localization of pea seed vicilin in ER and Golgi. *Protoplasma* **122**, 35-54
- Dauwalder, M., Whaley, W.G. (1982) Membrane assembly and secretion in higher plants. *J. Ultrastruct. Res.* **78**, 302-320
- Farquhar, M.G., Palade, G.E. (1981) The Golgi apparatus (complex) - (1954-1981) - from artifact to center stage. *J. Cell Biol.* **91**, 77s-103s
- Fernandez, D.E., Staehelin, L.A. (1985) Structural organization of ultrarapidly frozen barley aleurone cells actively involved in protein secretion. *Planta* **165**, 455-468
- Frens, G. (1973) Controlled nucleation for the regulation of the particle size in monodisperse gold suspensions. *Nature Phys. Sci.* **241**, 20-22
- Gardiner, M., Chrispeels, M.J. (1975) Involvement of the Golgi apparatus in the synthesis and secretion of hydroxyproline-rich cell wall glycoproteins. *Plant Physiol.* **55**, 536-541
- Gibson, R.A., Paleg, L.G. (1976) Purification of GA₃-induced lysosomes from wheat aleurone cells. *J. Cell Sci.* **22**, 413-425
- Greenwood, J.S., Chrispeels, M.J. (1985) Immunocytochemical localization of phaseolin and phytohemagglutinin in the endoplasmic reticulum and Golgi complex of developing bean cotyledons. *Planta* **164**, 295-302
- Herman, E.M., Shannon, L.M. (1984a) Immunocytochemical evidence for the involvement of Golgi apparatus in the deposition of seed lectin of *Bauhinia purpurea* (Leguminosae). *Protoplasma* **121**, 163-170
- Herman, E.M., Shannon, L.M. (1984b) Immunocytochemical localization of concanavalin A in developing Jack bean cotyledons. *Planta* **161**, 97-104
- Higgins, T.J., Jacobsen, J.V., Zwar, J.A. (1982) Gibberellic acid and abscisic acid modulate protein synthesis and mRNA levels in barley aleurone layers. *Plant Mol. Biol.* **1**, 191-215
- Jacobsen, J.V. (1984) Regulation of protein synthesis in aleurone cells by gibberellin and abscisic acid. In: *The biochemistry and physiology of gibberellins*, vol. 2. pp. 159-187, ed. Crozier, A., Praeger, New York
- Jacobsen, J.V., Beach, L.R. (1985) Control of transcription of

THIS PAGE BLANK (USPTO)

- α -amylase and rRNA genes in barley aleurone protoplasts by gibberellin and abscisic acid. *Nature* 316, 275-277
- Jacobsen, J.V., Higgins, T.J.V. (1982) Characterization of the α -amylases synthesized by aleurone layers of Himalaya barley in response to gibberellic acid. *Plant Physiol.* 70, 1647-1653
- Jones, R.L. (1969) Gibberellic acid and the fine structure of barley aleurone cells. II. Changes during the synthesis and secretion of α -amylase. *Planta* 88, 73-86
- Jones, R.L. (1972) Fractionation of the enzymes of the barley aleurone layer: evidence for a soluble mode of enzyme release. *Planta* 103, 95-109
- Jones, R.L., Jacobsen, J.V. (1982) The role of the endoplasmic reticulum in the synthesis and transport of α -amylase in barley aleurone layers. *Planta* 156, 421-432
- Kawasaki, I., Sato, S. (1979) Isolation of the Golgi apparatus from suspension cultured tobacco cells and preliminary observations on the intracellular transport of extensin-precursor. *Bot. Mag. (Tokyo)* 92, 305-314
- Kelly, R.B. (1985) Pathways of protein secretion in eukaryotes. *Science* 230, 25-32
- Kreil, G. (1981) Transfer of proteins across membranes. *Annu. Rev. Biochem.* 50, 317-348
- Locy, R., Kende, H. (1978) The mode of secretion of α -amylase in barley aleurone layers. *Planta* 143, 89-99
- Mitsui, T., Akazawa, T., Christeller, J.T., Tartakoff, A.M. (1985) Biosynthesis of rice seed α -amylase: two pathways of amylase secretion by the scutellum. *Arch. Biochem. Biophys.* 241, 315-328
- Palade, G. (1975) Intracellular aspects of the process of protein secretion. *Science* 189, 347-358
- Pyliotis, N.A., Ashford, A.E., Whitecross, M.I., Jacobsen, J.V. (1979) Localization of gibberellic acid-induced acid phosphatase activity in the endoplasmic reticulum of barley aleurone cells with the electron microscope. *Planta* 147, 134-140
- Rogers, J.C., Milliman, C. (1983) Isolation and sequence analysis of a barley α -amylase cDNA clone. *J. Biol. Chem.* 258, 8169-8174
- Roth, J., Bendayan, M., Orci, L. (1978) Ultrastructural localization of intracellular antigens by the use of protein A-gold complex. *J. Histochem. Cytochem.* 26, 1074-1081
- Stoward, P.J. (1985) News and views. *Histochem. J.* 17, 841-844
- Towbin, H., Staehelin, T., Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. USA* 76, 4350-4354
- Vigil, E.L., Ruddat, M. (1973) Effect of gibberellic acid and actinomycin D on the formation and distribution of rough endoplasmic reticulum in barley aleurone cells. *Plant Physiol.* 51, 549-558
- Vitale, A., Chrispeels, M.J. (1984) Transient N-acetylglucosamine in the biosynthesis of phytohemagglutinin: attachment in the Golgi apparatus and removal in protein bodies. *J. Cell Biol.* 99, 133-140
- Yomo, H., Varner, J.E. (1971) Hormonal control of a secretory tissue. *Curr. Top. in Devel. Biol.* 6, 111-144
- zur Nieden, U., Manteuffel, R., Weber, E., Neumann, D. (1984) Dictyosomes participate in the intracellular pathway of storage proteins in developing *Vicia faba* cotyledons. *Eur. J. Cell Biol.* 34, 9-17

Received 10 January; accepted 14 March 1986

THIS PAGE BLANK (USPTO)

Immunoelectron-microscopy localization of abscisic acid with colloidal gold on Lowicryl-embedded tissues of *Chenopodium polyspermum* L.

L. Sossountzov¹, B. Sotta¹, R. Maldiney¹, I. Sabbagh² and E. Miginiac¹

¹ Institut de Physiologie Végétale, CNRS, Laboratoire de Physiologie végétale, T 53, 5ème étage, Université P et M Curie, 4 Place Jussieu, F-75252 Paris Cedex 05, and

² Institut Pasteur, Domaine du Poirier, Lentilly, F-69210 l'Arbresles, France

Abstract. Further study on the localization of abscisic acid (ABA) has been undertaken at the ultra-structural level in *Chenopodium polyspermum* L. Axillary-bud-bearing nodes on the main axis were fixed with soluble 1-(3-dimethylaminopropyl)-3 ethyl carbodiimide, then postfixed with paraformaldehyde and embedded in Lowicryl K4M at -20°C . Ultrathin sections mounted on grids were successively incubated with rabbit anti-ABA antibodies and with gold-labelled goat anti-rabbit antibodies (40 nm particle size). Control sections treated with preimmune rabbit serum and ABA-preabsorbed antibodies were devoid of label. The background staining was very low with this technique. Quantitative analysis of the immunolabelling showed that two main sites of ABA accumulation could be defined: first, plastids in cortical cells and vascular parenchyma cells associated with sieve elements and xylem vessels; second, the cell cytoplasm and nucleus in the axillary bud tip and in procambial strands. In vascular bundles, the cambial cells showed no immunoreactivity. These observations support the hypothesis for the cytoplasmic synthesis of ABA which is subsequently trapped in plastids as cells mature.

Key words: Absciscic acid (localization) – *Chenopodium* – Immunogold electron microscopy – Lowicryl K4M.

Introduction

We recently initiated a series of experiments to investigate the localization of abscisic acid (ABA)

Abbreviations: ABA = abscisic acid; EDC = 1-(3-dimethylaminopropyl)-3 ethyl carbodiimide; GAR 40 = goat anti-rabbit antibodies labelled with colloidal gold of particle size 40 nm; IgG = immunoglobulin G

at the light-microscopic level using an indirect immunoperoxidase technique and fixed tissue embedded in paraffin or in Spurr's resin (Sotta et al. 1985). We found that, depending on the age of the cells, ABA is mainly localized in the meristematic cells of the slow-growing axillary bud tips and in plastids present in the parenchyma cells associated with sieve elements and xylem vessels in the vascular strands and in the inner cortex. Taking advantage of the recent finding that post-embedding of fixed tissues in the hydrophilic Lowicryl K4M resin and labelling with immunogold particles have a wide range of applications (De Mey 1983), we used this technique to provide further information on the precise location of ABA at the electron-microscope level. We confirm here that ABA is mainly trapped in plastids of the more differentiated parenchyma cells associated with phloem and xylem tissue and in the inner cortical cells, and more precisely show that in the early stages of tissue differentiation a preferential localization of antibodies is seen over the cytoplasm and the nuclei.

Material and methods

Plant material and electron microscopy. Young axillary nodes beneath the terminal bud, the second nodes bearing two axillary buds, and cotyledonary nodes and their associated buds were collected from normally watered plants of *Chenopodium polyspermum* L. (see Fig. 1 in Sotta et al. 1985) and treated as previously described. In brief, they were rapidly dropped in a 2% aqueous solution of water-soluble 1-(3-dimethylaminopropyl)-3 ethyl carbodiimide (EDC), an essential tissue processing which allows fixation of ABA in situ. As we have recently shown, the omission of this prefixation step led to the absence of immunolabelling (see Fig. 4 in Sotta et al. 1985). The organs were then postfixed in buffered paraformaldehyde solution. After an extensive rinse in phosphate-buffered saline, they were rapidly dehydrated in an ethanol series. After the 90% ethanol rinse, they were infiltrated with Lowicryl K4M (Balzers Union, Paris,



THIS PAGE BLANK (USPTO)

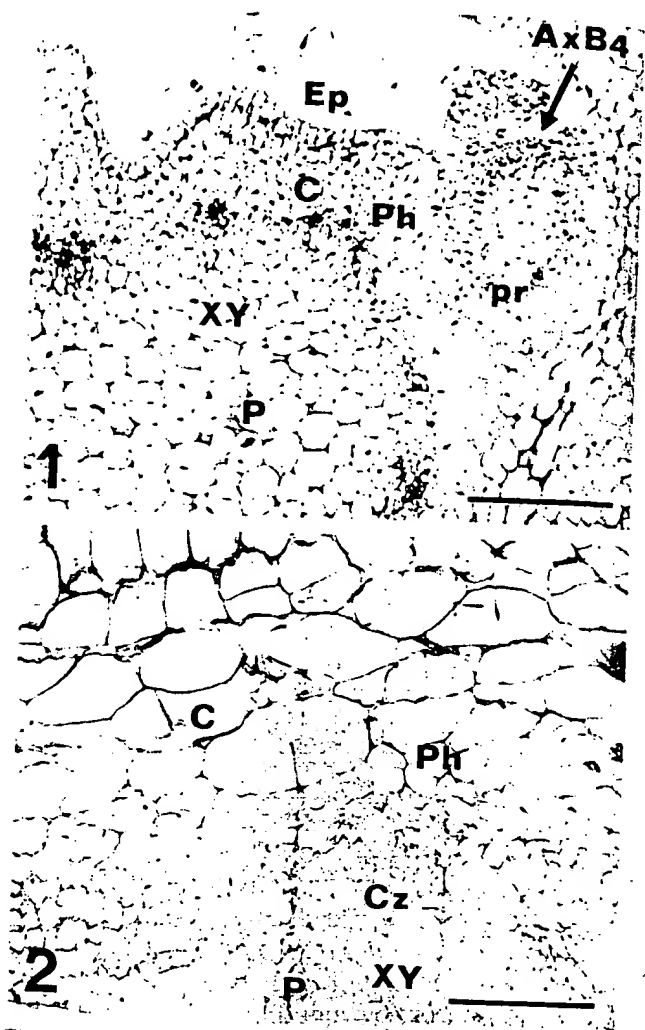
France). Infiltration and polymerisation by UV irradiation for 48 h were carried out according to the technique of Roth et al. (1981) and Carlemalm et al. (1982) and were performed at -20°C . Polymerisation by UV irradiation was continued at room temperature for a further 2 d. Ultrathin sections were cut with glass knives on a Reichert OMU2 microtome (Reichert, Wien, Austria) and picked up on 150-mesh copper or nickel grids with carbon-coated 0.3% formvar films. With regard to the difficulty of cutting encountered by ourselves and by other plant cytologists (Craig and Goodchild 1982; Herman and Shannon 1984), only entire gold or silver sections were mounted on grids, whatever their thickness. This had no influence on the labelling because only the section surface was labelled, as demonstrated by Bendayan (1984). Adjacent semi-thin sections, stained with Ehrlich's Hematoxylin, were used in an attempt to localize precisely the tissular organization of the anatomically heterogeneous nodes. Sections were cut to obtain meristematic cells of the axillary bud apex and associated leaf primordia, vascular bundles of the stem and parenchyma cells associated with the vascular area as seen in Figs. 1 and 2.

Immunocytochemical procedure. Specific antibodies against ABA were produced in rabbits and characterized as described previously (Sotta et al. 1985; Leroux et al. 1985; Maldiney et al. 1986). Furthermore, enzyme-linked immunosorbent assay (ELISA) measurements have shown no cross-reactivity with phaseic and dihydrophaseic acids (samples given by Dr. Mappelli, CNR, Pisa, Italy). All immunoelectron-microscopy staining procedures were carried out at room temperature. We employed here a two-step immunocytochemical procedure. All antibodies were diluted with 0.1 M 2-amino-2-(hydroxymethyl)-1,3-propanediol (Tris)-HCl buffer, with 0.1 M NaCl (TBS), pH 7.6, containing normal goat serum ($1\text{ mg}\cdot\text{ml}^{-1}$), 0.1% Lambda Carrageenan gelatin (Type IV; Sigma St. Louis, Mo., USA), and 0.1% Triton X-100 to prevent non-specific binding of antibodies to the tissues. All immunoreagents and buffer washes were freshly Millipore-filtered ($0.22\text{ }\mu\text{m}$ pore size). Grid-mounted sections were first floated, section-side down, for 15 min on a drop of buffered normal goat serum. Without rinsing, they were then transferred onto drops of anti-ABA antibodies for 1 or 2 h, using a dilution of $250\text{ }\mu\text{g}\cdot\text{ml}^{-1}$. After incubation, they were washed in several changes of TBS and floated for 10 min on a 10-fold diluted solution of colloidal-gold (40 nm particle size)-labelled goat anti-rabbit immunoglobulin G (IgG) (GAR 40; Janssen Pharmaceutica, Beerse, Belgium). After rinses with TBS, the grids were post-stained with saturated uranyl acetate in distilled water for 10 min and in alkaline lead citrate for 5 min. The counterstaining did not mask the electron-dense gold particles but partially compensated for the omission of OsO_4 during the fixation step.

Different controls were employed to establish the specificity of the immunogold procedure: i) incubation with preimmune rabbit IgG or normal goat IgG instead of anti-ABA IgG; ii) incubation with anti-ABA antibodies previously adsorbed with an excess of ABA for 24 h at 4°C ($0.1\text{ }\mu\text{M}$ *cis*(\pm)ABA $\cdot\text{ml}^{-1}$); iii) incubation with diluted GAR 40 only, omitting the first antibody step. Some cellular compartments lacked labelling and served as internal controls.

Electron micrographs were taken with a Hitachi HU 11 B electron microscope at calibrated magnifications and at an acceleration voltage of 75 kV.

Brief description of the bud-bearing nodes. The tissue arrangement at different levels of the main axis is shown in Figs. 1 and 2 and its terminology follows that developed in our preceding report (Sotta et al. 1985). Thin sections through the node

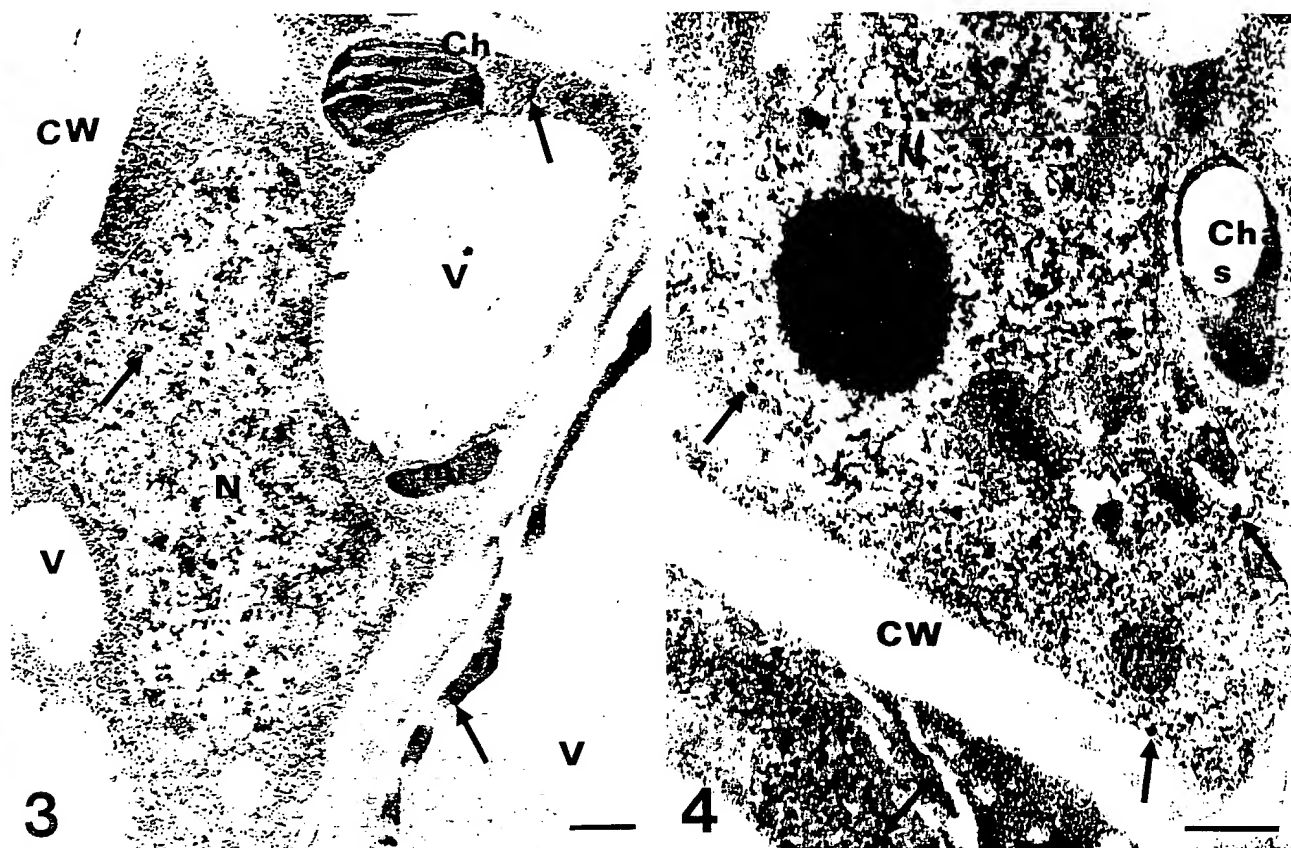


Figs. 1-2. Structural arrangement of the tissues in nodes at two levels of *Chenopodium* plants, embedded in Lowicryl K4M, thick-sectioned and Erlich's Hematoxylin-stained. Fig. 1. Section through the node 4 and its associated axillary bud A x B₄ at the plant tip. The anatomical arrangement is similar in the second node, at the median part of the plant. Fig. 2. Section through the cotyledonary node; nuclei appear black. Notice that in the old cotyledonary node, the nuclei are rarely distributed. Bar = 50 μm ; $\times 190$.

Abbreviations for Figs. 1-15: G=cortex; Ch=chloroplast; Cha=chloroamyloplast; CW=cell wall; CZ=cambial zone; mi=mitochondrion; V and Nu=nucleus and nucleolus; P=pith; pr=procambium, linking the axillary bud to the main axis at node; Ph and XY=phloem side and xylem side of the vascular bundles present in the stem at node; s=starch grain; v=vacuole.

contain: i) the nearly quiescent meristematic cells of the bud apex and its associated leaf primordia; ii) beneath the apex, the bud branch with procambial cells in a non-differentiated state; and iii) in the main axis, a zone comprising vascular strands with phloem, xylem and cambial zone, surrounded by parenchyma cells of the cortex and pith. For the cotyledonary node (Fig. 2), we examined only the cells of the main axis because the associated buds in the three blocks we cut were poorly impregnated by the acrylic resin, leading to bad quality sections.

THIS PAGE BLANK (USPTO)



igs. 3-4. Control and reaction specificity. Thin sections through meristematic cells from the axillary bud 4 were treated with preimmune rabbit serum (Fig. 3) or with ABA-preadsorbed anti-ABA antibodies (Fig. 4) followed by the IgG-gold complex. Very few (arrows) gold particles, isolated or associated in pairs, are present over the different cell structures. Structural preservation is reasonably good. Note, however, the negative contrast of some membranes. Plasmalemma and tonoplast are not visible. Bar = 0.5 μm ; $\times 16500$ and 23500 , respectively

Particular ABA-containing cells were selected for observation at the electron-microscope level on the basis of our previous study (Sotta et al. 1985). In each category, at least 10 sections from three sets of nodes were examined.

Quantitative analysis. Sections in selected areas were photographed at a magnification of $\times 10000$ and $\times 16000$. Printing of the negatives was standardized, and they were enlarged $\times 3$. Areas on the micrographs corresponding to cytoplasm, plastids and nuclei, including nucleoli, were measured by placing a transparent graph paper over each enlarged photograph. The number of gold particles was counted manually. Countings were done from 5-12 electron micrographs corresponding to the same number of cells if in a meristematic state and to a higher number for differentiated cells. We have actually to precise that in the fully differentiated or senescent tissues of the stem, i.e. at the median and basal levels of the plant, cells are very large and highly vacuolated. Moreover, there is a very low frequency of nuclei in the thin sections (Fig. 2) so that only a few values were obtained. Therefore, countings are given for information only. Furthermore, as cells enlarge, the starch grains became very large structures, occupying one- or two-thirds of the plastid stromal space. Therefore, the number of gold particles was counted over the entire plastid or over its stromal subcompartment.

The labelling density was expressed as the number of gold particles per μm^2 . In order to reduce the effect of varying exper-

imental conditions on the final calculations, a set of three different axillary-bud-bearing nodes was processed at the same time.

Results

Specificity of immunogold staining

Controls showed very low levels of a random distribution of gold particles over sections and grid support films. Control sections with an essential reactant omitted gave appropriately negative reactions. The serial-section analysis revealed that only slight non-specific staining was observed when rabbit preimmune serum took the place of the primary anti-ABA serum, whatever the cells observed, in a mature or meristematic state (Fig. 3). The labelling reaction generated by the anti-ABA antibodies was inhibited by preincubation with the corresponding antigen in excess (pure *cis*(\pm)ABA, Figs. 4, 11, 13). All these controls produced less than one gold particle per μm^2 (Table 1). Furthermore, when this very low non-specific staining was

THIS PAGE BLANK (USPTO)

Table 1. Control experiments showing labelling densities over the cellular compartments at three different levels of plants of *Chenopodium polyspermum*. Values are expressed as the mean number of gold particles per $\mu\text{m}^2 \pm \text{SEM}$; n = the number of areas over which the values were obtained. *Nd 4* = node 4 near the terminal bud; *Nd 2* = node 2 in the median part of the plant; *Cot Nd* = cotyledonary node

Controls	Normal rabbit IgG	Preadsorbed anti-ABA IgG
<i>Nd 4</i>		
cytoplasm	0.3 ± 0.1 ($n=6$)	0.2 ± 0.1 ($n=7$)
nucleus	0.2 ± 0.1 ($n=5$)	0.4 ± 0.2 ($n=6$)
plastids	0.3 ± 0.1 ($n=6$)	0.3 ± 0.2 ($n=6$)
<i>Nd 2</i>		
cytoplasm	0.3 ± 0.2 ($n=8$)	0.2 ± 0.1 ($n=7$)
nucleus	1.2 ± 0.4 ($n=5$)	0.5 ± 0.2 ($n=6$)
plastids	0.7 ± 0.3 ($n=8$)	0.3 ± 0.1 ($n=7$)
<i>Cot Nd</i>		
cytoplasm	0.2 ± 0.1 ($n=11$)	—
nucleus	0.9 ± 0.3 ($n=5$)	—
plastids	0.3 ± 0.2 ($n=9$)	—

observed, gold particles were more often seen isolated (Figs. 3, 11, 13), rarely in pairs (Fig. 4) but never as clusters. Only GAR 40 showed no binding to tissue structures when applied directly to thin sections.

Immunocytochemical labelling

i) *In the axillary buds.* Figure 5 shows a portion of a meristematic cell from the young bud tip. In these small cells with their large nuclei and in which plastid starch grains were nearly absent, the gold labelling was essentially cytoplasmic and nuclear. When present, plastids with or without starch were poorly labelled (Fig. 5). In the nucleus, gold particles were also associated with the electron-dense nucleolus. Occasionally, they were close to the surface of the vacuole. Mitochondria (not present in Fig. 5) had a very low level of labelling equal to background. In the leaf primordia, the gold label was also distributed throughout the cytoplasm and in the nuclei but remained slight over plastids. Most of the gold particles appeared isolated, but clusters were not rare in these cells of the apical tip. The largest aggregates were mainly associated with the nucleoplasm. Therefore, they cannot be considered as an artifact if compared with the controls.

In the underlying layers, procambial cells and closely associated cells of the inner cortex showed a greater morphological differentiation (Figs. 6, 7). Plastids were larger, contained some starch grains, and were more immunoreactive. Mitochondria

Table 2. Effects of axillary-bud maturation of *Chenopodium* on the immunogold labelling of anti-ABA IgG over cytoplasm and plastids in apices and procambium (*pr*) strands. Values are expressed as the mean number of gold particles per $\mu\text{m}^2 \pm \text{SEM}$; n = number of observations. $A \times B_4$ and $A \times B_2$ = in axillary buds 4 and 2; *pr* = procambium. Cotyledonary buds were not studied

Levels on the plant	Cytoplasm	Plastids ^a	Stroma ^a
<i>Ax B4</i>			
apex	8.6 ± 1.2 ($n=8$)	2.4 ± 0.2 ($n=7$)	2.4 ± 0.2 ($n=7$)
pr	7.9 ± 1.2 ($n=5$)	3.2 ± 0.9 ($n=5$)	3.2 ± 0.9 ($n=5$)
<i>Ax B2</i>			
apex	2.8 ± 0.8 ($n=7$)	3.9 ± 0.4 ($n=5$)	7.4 ± 0.8 ($n=5$)
pr	3.3 ± 0.2 ($n=4$)	6.8 ± 0.3 ($n=3$)	13.5 ± 1.7 ($n=3$)

^a Labelling was calculated individually for each starchy plastid over the entire organelle and over its stromal space

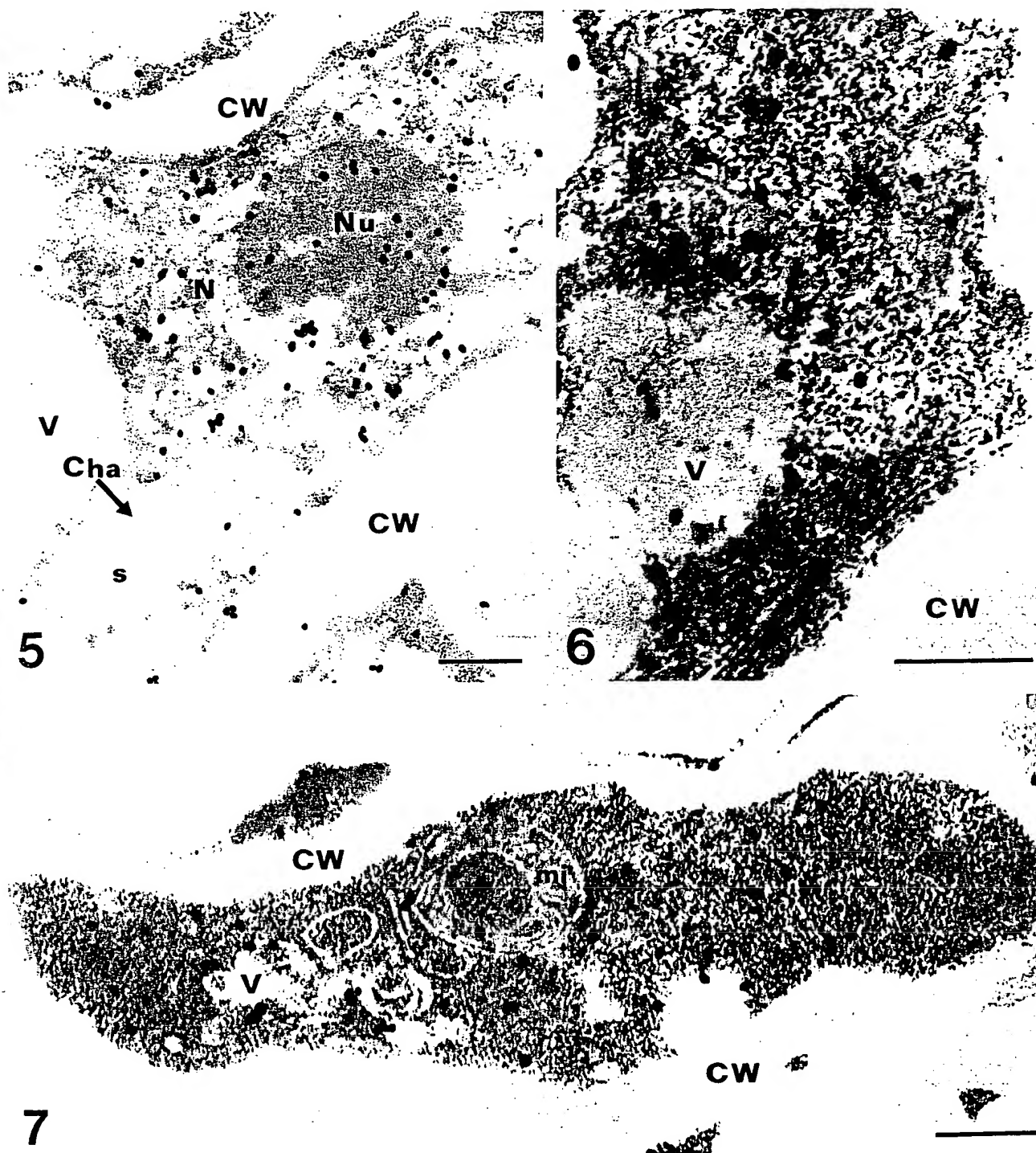
Table 3. Ratio of plastid/cytoplasm distribution of immunogold labelling after anti-ABA treatment in axillary-bud-bearing nodes of *Chenopodium* during their differentiation. The mean distribution ratios were calculated from the number of gold particles over entire plastids and over their stromata (number in parentheses). For abbreviations, see Table 1

Levels on the plant	Apex	Procambium	Inner cortex	Parenchymal cells in	
				phloem	xylem
<i>Nd4</i>	0.3	0.4	2.3 (3)	1.2 (1.3)	—
<i>Nd2</i>	1.4 (2.6)	2.1 (4.1)	5.5 (9.7)	2 (3.9)	2.6 (6.6)
<i>Cot Nd</i>	—	—	2.8 (5.6)	2.5 (4.7)	8.5 (13.8)

were poorly labelled (Fig. 6), but the label remained distributed throughout the cytosol (Figs. 6, 7) and nuclei.

Table 2 shows that there was a twofold reduction in labelling over the cytoplasm as the bud aged and a twofold increase in gold particles over plastids. This fact was more evident when the density of gold particles over plastids was calculated on a stromal basis, minus the starch areas which became larger in the bud branch beneath the apical tip but were very rarely labelled. As seen in Table 3, the mean ratios of the distribution of labelling between plastids and cytoplasm became greater than 1 as the cells matured and the buds aged. There was a twofold increase of these ratios during differentiation. During the same time, nuclear staining remained relatively constant and showed only a downward trend as the bud matured. Here again, the number of gold particles overlying mitochondria, starch grains, vacuoles and cell walls was negligible and showed the high degree of specificity of the immunogold method.

THIS PAGE BLANK (USPTO)



Figs. 5–7. Immunogold localization of ABA in axillary bud 4. The thin sections were treated with anti-ABA rabbit serum ($250 \mu\text{g} \cdot \text{ml}^{-1}$) followed by goat anti-rabbit gold-labelled IgG. Fig. 5. In a meristematic cell of the bud apex, most of the gold particles are located over the cytosolic compartments and over the nucleus (some of which have suffered from compression), leaving the cell walls, vacuoles and chloroamyloplasts unlabelled. Note some clusters of gold particles especially over the nucleoplasm. Bar = $0.5 \mu\text{m}$; $\times 26\,500$. Figs. 6, 7. Immunogold staining of differentiating cells in a grazing section of the inner cortex (Fig. 6) and in a procambial cell (Fig. 7), showing a cytosolic distribution of ABA while cell walls and vacuoles are without specific label. Note the absence of plastids in these two thin sections. Bar = $0.5 \mu\text{m}$; $\times 45\,000$ and $33\,000$, respectively.

THIS PAGE BLANK (USPTO)

ii) In
analy:
outer
in me
cambi
Fc
came
co int
and se
Outer
any g
sieve e

Table 4.
buds o
pressed
n the
w low
tion onl

Levels o
the plan

Nd4

Nd2

(Nd

Table 5.
plants.
values a

Levels o

Nd4

cytopl
plastic
stroma

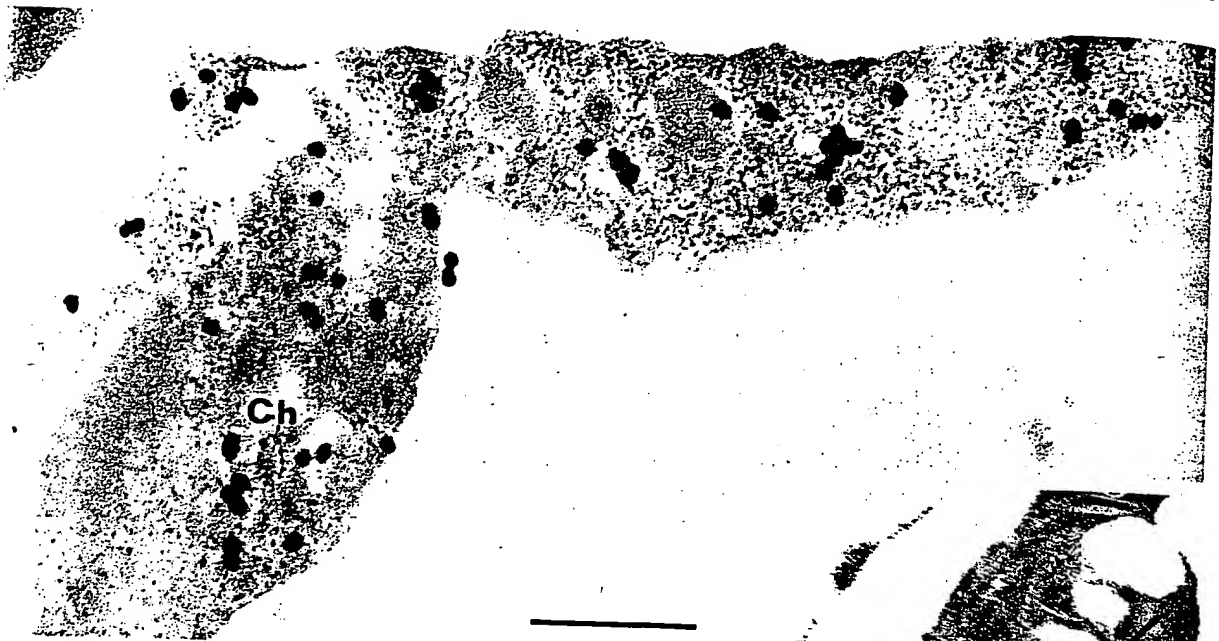
Nd2

cytopl
plastic
stroma

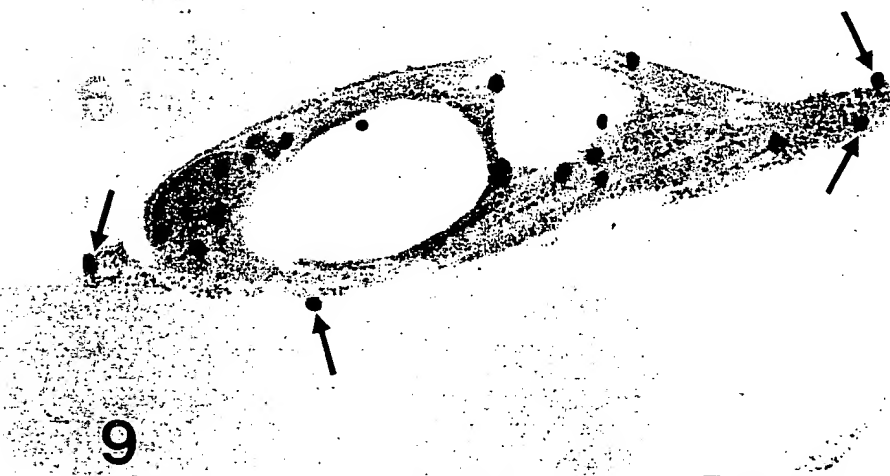
Cot Nd
cytopl
plastic
stroma

Figs. 8-1
the you
the med
are seen
node, a
absence
× 65000

Fig. 11.
labelli



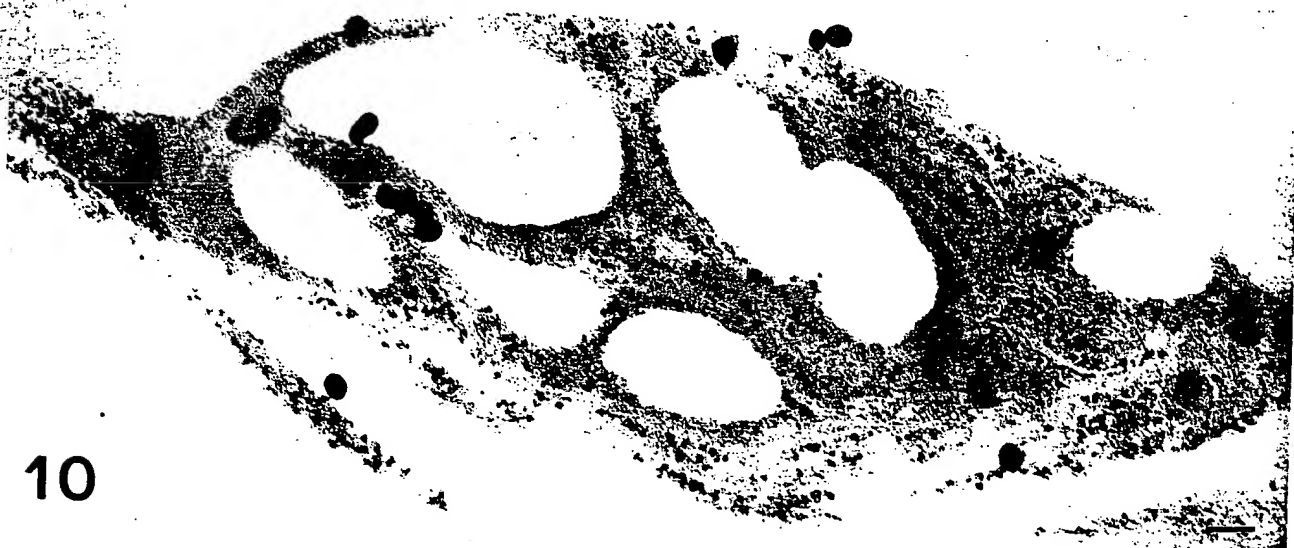
8



9



11



10

THIS PAGE BLANK (USPTO)

ii) *In the stem.* The observations and quantitative analysis were made in the cortex, where inner and outer cortical cells were analyzed separately, and in medullary cells. In the vascular bundles, phloem, cambial and xylem-parenchyma cells were studied.

For plastids, as cells enlarged, starch grains became larger (Figs. 8, 9, 10) so that gold-grain counts were always estimated over their total area and separately over their living subcompartments. Outer cortical and medullary cells were devoid of any gold conjugates. Xylem lignified vessels and sieve elements, decorated with the characteristic P-

protein network and P-protein bodies (Fig. 15), showed less than one gold particle per μm^2 and served as internal controls. Interestingly, cambial cells (Fig. 14) showed a labelling slightly higher than background over the cytoplasm and plastids in the mature buds but within the limits of specificity. By contrast, the nuclear labelling was not specific (Table 4). The level of label in cambial cells was less than $1 \mu\text{m}^{-2}$ and represents the measured average of 16 such fields from the three sets of buds analysed in this study. Three major types of labelled cells were readily recognized along the main axis. Immunoreactivity which was ABA-specific (Fig. 11) was predominantly localized over chloroamyloplasts of the inner cortex and in the phloem and xylem parenchyma cells. No specific label was found over mitochondria (Fig. 8). In these differentiated cells with their large vacuoles, plastids were closely pressed to the non-labelled cell wall. The cytoplasm appeared as a thin layer along the wall (Fig. 8).

The density of gold over each plastid varied from one to another, so that in some cases the standard errors of the means appeared high. Nevertheless, compared with the cytosolic labelling at the level of the meristem (Table 2), there was a twofold decrease of the cytosolic immunoreactivity in the nodal cells at the tip of the plant. Cytoplasm

Table 4. Nuclear immunogold labelling in nodes and in axillary buds of *Chenopodium* during differentiation. Values are expressed as the mean number of gold particles per $\mu\text{m}^2 \pm \text{SE}$; n = the number of observations. For some tissues, this number is low and the nuclear labelling was calculated for information only. For abbreviations, see Table 1

Levels on the plant	Apex	Procambium	Differentiated cells	Cambial cells
Nd4	10.8 ± 1.0 ($n=6$)	8.8 ± 2.0 ($n=4$)	10.8 ± 2.0 ($n=4$)	0.6 ($n=2$)
Nd2	7.6 ± 0.9 ($n=4$)	7.0 ($n=2$)	7.1 ± 2.4 ($n=4$)	0.4 ± 0.3 ($n=3$)
Cot Nd	-	-	8.5 ± 2.2 ($n=6$)	0.7 ± 0.4 ($n=4$)

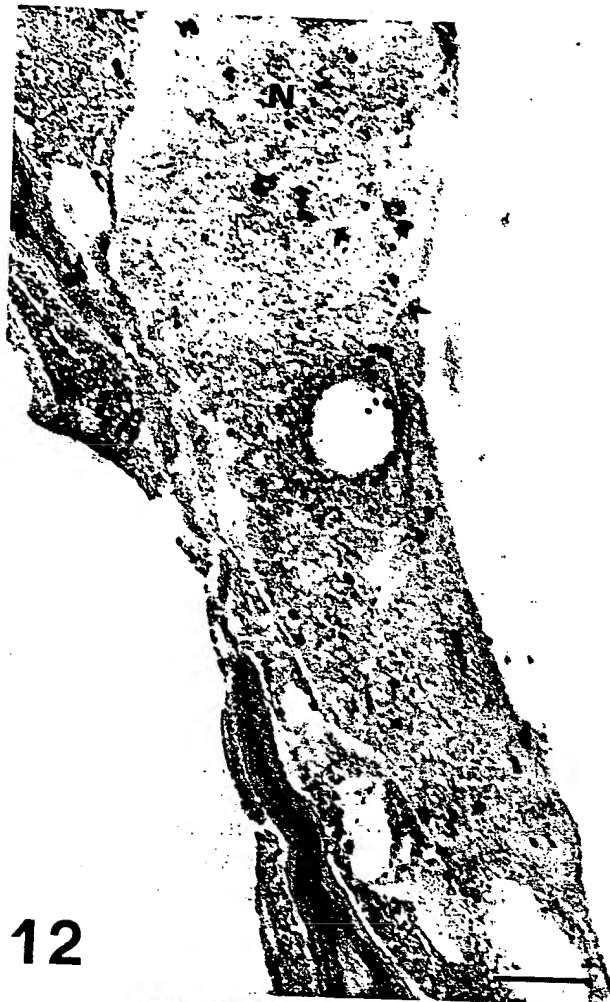
Table 5. Immunogold labelling after anti-ABA treatment of differentiated cells in the nodes at different levels of *Chenopodium* plants. Values are expressed as the mean number of gold particles per $\mu\text{m}^2 \pm \text{SEM}$; n = the number of areas over which the values are obtained. For abbreviations, see Table 1

Levels on the plant	Inner cortex	Phloem parenchyma cells	Xylem parenchyma cells	Cambial cells
Nd4				
cytoplasm	4.5 ± 1.1 ($n=6$)	6.0 ± 1.1 ($n=6$)	-	0.2 ± 0.1 ($n=5$)
plastids	10.3 ± 0.7 ($n=6$)	6.9 ± 1.2 ($n=6$)	-	0.3 ± 0.2 ($n=5$)
stroma	13.7 ± 1.8 ($n=6$)	7.8 ± 1.0 ($n=6$)	-	0.7 ± 0.4 ($n=5$)
Nd2				
cytoplasm	1.1 ± 0.5 ($n=10$)	2.0 ± 0.4 ($n=4$)	1.1 ± 0.4 ($n=6$)	1.0 ± 0.4 ($n=3$)
plastids	6.1 ± 0.5 ($n=10$)	3.9 ± 0.5 ($n=3$)	2.9 ± 0.5 ($n=5$)	0.9 ± 0.3 ($n=3$)
stroma	10.7 ± 0.7 ($n=10$)	7.8 ± 1.1 ($n=3$)	7.2 ± 1.3 ($n=5$)	1.2 ± 0.5 ($n=3$)
Cot Nd				
cytoplasm	1.9 ± 0.4 ($n=4$)	2.7 ± 1.6 ($n=4$)	1.1 ± 1.0 ($n=4$)	1.2 ± 0.7 ($n=11$)
plastids	5.4 ± 1.4 ($n=5$)	6.7 ± 2.0 ($n=4$)	9.3 ± 3.7 ($n=4$)	0.7 ± 0.5 ($n=8$)
stroma	10.6 ± 2.0 ($n=5$)	12.7 ± 3.9 ($n=4$)	15.2 ± 4.0 ($n=4$)	1.1 ± 0.5 ($n=8$)

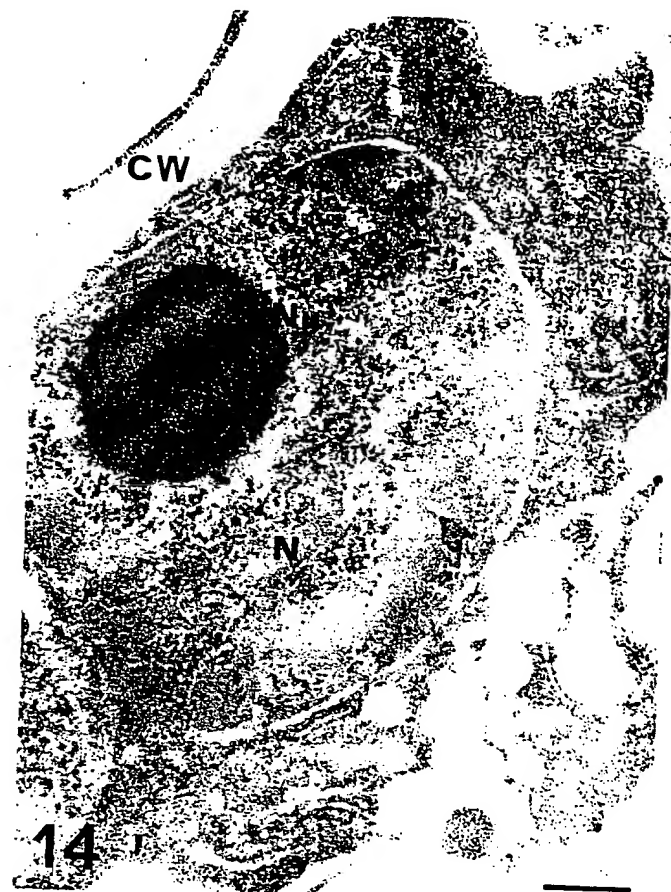
Figs. 8–10. Immunogold localization of ABA in the inner cortex of the nodes at different levels of the main stem. **Fig. 8.** In the younger node (Nd4), gold particles are located over the chloroplast and the cytoplasm. Bar = $0.5 \mu\text{m}$; $\times 43000$. **Fig. 9.** In the median node (Nd2), the labelling was mainly chloroplastic but starch grains are free of label. Some gold particles (arrows) are seen in the cytoplasm and close to the non-labelled cell wall and vacuole. Bar = $0.5 \mu\text{m}$; $\times 41000$. **Fig. 10.** In the cotyledonary node, a starchy plastid at a higher magnification shows that gold particles are distributed over the stromal space. Note the absence of label on starch grains and the negative contrast of the plastid envelope and its thylakoid membranes. Bar = $0.1 \mu\text{m}$; $\times 65000$.

Fig. 11. In the same node, an ultrathin section stained with anti-ABA antibodies preadsorbed with ABA shows the specificity of labelling: there are only two gold particles (arrows) over the starchy plastid. Bar = $0.5 \mu\text{m}$; $\times 22000$.

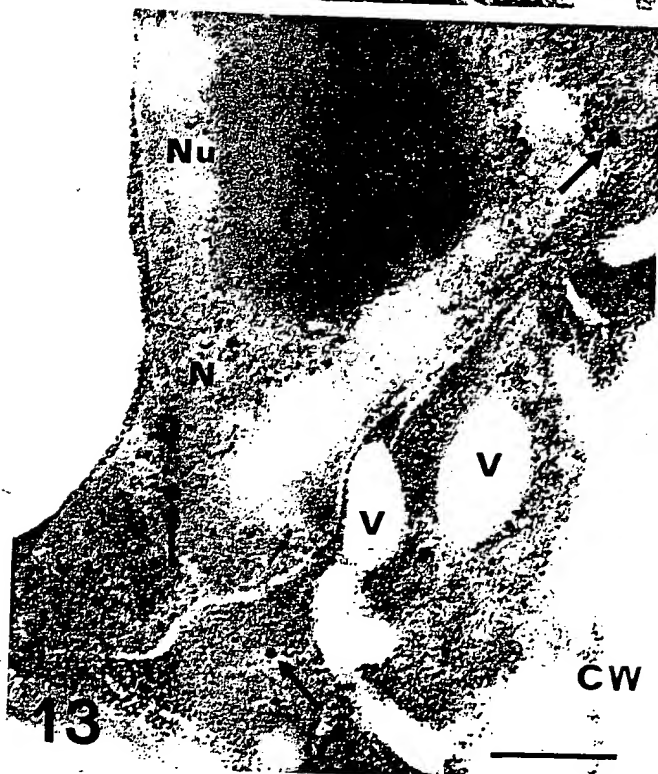
THIS PAGE BLANK (USPTO)



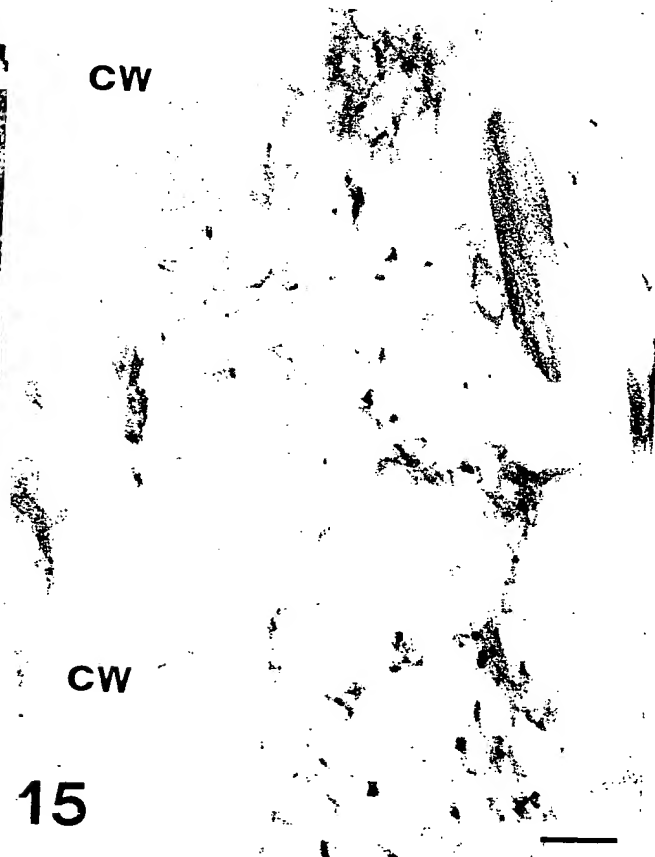
12



14



13



CW

15

labell
2 and
slight
ligh-
gold
prese
cent t
lopes
10. C
corte
paren
at the
appea
previ
the ra
and c
ble 3)
cr. use
stem (

At
during
ther i
apices
phasiz
ferent
not al
st. fici
when
young
labelli
an exc
particl
areas (

Discus

Immu
tion w
terest
ically
nodes.
is syn
startec
tion or
ration
nous A

←
Figs. 12-
at the r
It shows
25000 a
section
22000
19500



THIS PAGE BLANK (USPTO)

labelling decreased abruptly at the level of node 2 and remained constant downward (Table 5). This slight immunoreactivity was not detected at the light-microscope level (Sotta et al. 1985). When gold particles, isolated or associated in pairs, were present in the cytosol, most of them occurred adjacent to the outer surface of plastids, the outer envelopes of which were not clearly resolved (Figs. 9, 10). Gold clusters were rarely observed. The inner cortex became more immunolabelled than other parenchyma cells in the vascular bundles, except at the cotyledonary level where xylem parenchyma appeared the most immunoreactive, as we have previously noted (Sotta et al. 1985). Consequently, the ratios of labelling between chloroamyloplasts and cytoplasm always became greater than 1 (Table 3). In the stromal space, there was a great increase of the immunoreactivity at all levels of the stem (Table 5).

Another fact to be noted is the maintenance during cell maturation of the nuclear labelling either in the more meristematic cells of the bud apices or in the procambial cells. As we have emphasized in Material and methods, the highly differentiated state of the cells in the main axis did not allow us to estimate the gold labelling on a sufficient number of nuclei, but, in almost all cases, when observed, they were labelled as in the younger cells of the axillary buds (Fig. 12) and this labelling was specific (Fig. 13). Cambial cells were an exception since an average of less than one gold particle per μm^2 could be measured over nuclear areas (Fig. 14).

Discussion

Immunocytological techniques used in combination with electron microscopy are of particular interest since they enable us to distinguish in anatomically complex structures, such as bud-bearing nodes, the cellular compartments in which ABA is synthesized or trapped. This investigation was started with a view to obtaining further information on ABA distribution in cells during their maturation. We have previously shown that endogenous ABA can be effectively fixed by EDC on the

cellular protein network and detected afterwards by an immunoenzymatic technique at the tissular level (Sotta et al. 1985). Numerous controls were carried out on sections and on a model system. The validity of the technique was reinforced by the correlation we have presented between the increase in ABA levels and immunoreactivity in water-stressed plants compared with normally watered plants. Furthermore, as apices were heavily immunostained, it was unlikely that a possible gradient of penetration of EDC had occurred. Here we show that the use of an on-grid immunogold labelling procedure, introduced by Roth et al. (1978), is effective on ultrathin sections of plant tissues after low-temperature methacrylate embedding, and allows comparative quantitation of gold particles. The Lowicryl K4M embedding technique is not usually employed by plant cytologists to reveal antigenicity (Craig and Goodchild 1982; Herman and Shannon 1984; Tomenius et al. 1983). Only one report is presently available on the localization of plant growth regulators at the electron-microscope level using an epoxy resin (Zavala and Brandon 1983). Although the structural preservation of the tissue was not as optimal as with animal tissues, it was nonetheless sufficient to localize an antigen and render possible the observation of some fine cellular details. Plastids, mitochondria and nuclei were well-enough preserved but membranes were not well defined as compared with more conventional electron-microscope techniques. One also notices that it was difficult to obtain good sections in lignified tissues and the impregnation of old meristematic cells was poor. The material embedded in Lowicryl K4M showed a negative contrast for membranes, as recently pointed out by Goodchild et al. (1985a) and Shaw and Henwood (1985), leading to difficulties in distinguishing, for instance, the chloroplast envelope from the adjacent cytoplasm, but thylakoids were visible in more detail. The mitochondrial envelope and inner membranes were poorly preserved and the plasma membranes and tonoplast were not visible. However, subcellular details in other compartments were remarkably well delineated despite the omission of OsO_4 postfixation and the mild fixation with paraformaldehyde following the neces-

Figs. 12–15. Immunogold labelling and control specificity in phloem and cambial cells from vascular bundles of the main stem at the median level (Nd2). Figs. 12–13. In a mature phloem parenchymal cell, the nucleus is more labelled than the cytosol. It shows some gold clusters. This labelling is specific as shown by the preadsorbed anti-ABA serum treatment (Fig. 13). Bar = 0.5 μm ; $\times 25000$ and $\times 33000$. Fig. 14. In a cambial cell, by contrast, the cytosolic and nuclear labelling was very low when the thin section was incubated with anti-ABA serum followed by GAR-40 complex. Gold particles are isolated or in pairs. Bar = 0.5 μm ; $\times 22000$. Fig. 15. In a sieve element, there is no labelling on the P-protein network and the P-protein bodies. Bar = 0.5 μm ; $\times 19500$.

THIS PAGE BLANK (USPTO)

sary carbodiimide-prefixation step (Yamamoto and Yasuda 1977). Cytoplasmic ribosomes were clearly resolved and nucleoli appeared well stained. These inadequacies have led some animal and plant physiologists to use other acrylic monomers, less complex than Lowicryl K4M, which lead to a positive immunocytochemical reaction after short periods in the primary antiserum and good membrane contrast, thus providing more precise localization of gold labelling in these structures. These LR white or LR gold resins (Craig and Miller 1984; Goodchild et al. 1985a, b; Greenwood and Crispeels 1985; Harris and Croy 1985; Newman et al. 1983) should be tested in future investigations. However, we show here with Lowicryl K4M that tissue preservation is appreciably better in the more meristematic cells of the bud apex than in the highly vacuolated cells of the main axis from which sections suffered from compression (Craig and Miller 1984).

Nevertheless, the two-step postembedding staining of thin sections with IgG-Gold labelling strengthens the findings of our previous report (Sotta et al. 1985) in which ABA was located by means of indirect peroxidase-antiperoxidase staining using paraffin sections and semi-thin etched resin sections. It also provides new information since it allows a comparative quantitation of labelling at the subcellular level. The control experiments and the internal controls we encountered clearly show that the labelling is not an artifact arising from the non-specific binding of antibodies to the tissue sections. Positive cells were scattered among negative cells and a good number of equally aged cells by experiment gave the same labelling. No appreciable labelling was found associated with cell walls, mitochondria, vacuoles, and P-proteins in the sieve tubes and starch grains. There was an almost complete absence of gold particles in the outer cortical cells, and in cambial and medullary cells. Preadsorbed ABA antibodies, which constitute an essential control for determining the specificity of the immunolabelling, produced less than one gold particle per μm^2 over cytoplasm, nuclei and plastids.

Comparing the location of ABA in axillary-bud-bearing nodes of different ages, we demonstrate here unambiguously that ABA is heterogeneously distributed within these organs. Age-dependant specific cell layers and specific organelles contain the plant growth regulator. In mature tissues of the main axis, ABA is mainly sequestered into chloroamyloplasts present in vascular parenchyma cells and cortical cells at the immediate vicinity of the phloem bundles. Similar labelling was

not observed in other cell types: medullary and external cortical cells, epidermal cells, xylem vessels and cambial cells. In younger cells, in a non-specialized state, belonging to the apical tip and its newly formed leaf primordia, as in the underlying procambial strands, ABA is found in the cytoplasm and in the nucleus. Plastids show here poor labelling. Cytosolic ABA declines with time in older buds. In the younger apices, it was repeatedly found that 75% and 25% of the immunoreactivity occurred within the cytoplasm and plastids respectively. At later stages of differentiation, these values were reversed. The marked decrease in cytoplasmic immunoreactivity was associated with a marked increase in plastid immunoreactivity. As cells enlarge, the increase in size of the amyloplasts parallels the development of the starch grains, so that the number of gold particles found in the stromal space provides a better estimation. These latter data indicate that the cytosolic localization of ABA declines with time in older buds, and plastids clearly increase their ability to accumulate or synthesize ABA. This heterogeneous age-dependant distribution of ABA may be related to different cellular sites of ABA synthesis or sequestration, in agreement with the biochemical observations and experiments with radioactive ABA. As suggested by previous workers, chloroplasts would be the site of ABA synthesis (Milborrow 1979; Zeevaert 1977). This view was criticized by Walton (1980). According to Hartung et al. (1981), 90% of the total ABA present in leaves is located within the chloroplasts after being biosynthesized in cytoplasm. Our study provides ultrastructural evidence to support the concept of two ABA accumulation sites depending on the age of the cells. It may be related to the possible ABA synthesis in cytoplasm and further trapping in plastids as the cells mature (Cowan et al. 1982).

The present data show that considerable antigenicity within the meristematic cells of the bud, whatever its age, and within some mature cells of the stem, except cambial cells, is also associated with the nucleus. The nuclear staining is evident in this study. In our previous report, only a slight immunoreactivity was found in the nuclei of the bud apex. A comparison of the results obtained in these two studies indicates that the immunoreactivity may be related to the methods used. The procedure using Lowicryl K4M at low temperature, compared with that used previously (conventional embedding in paraffin or in Spurr's resin at high temperature), seems to be a better way to preserve the antigenicity of some cellular compartments. The meaning of the nuclear localization

L. So
is at
lishe
obse
the
lecul
stoo
certa
re
level
Th
mecl
diffe
elect
mati
ucts.
in di
for s
whe
acce
pear
of o
are p
We a
to Dr
micro

Refs

Bend
m
J.
Carle
op
di
Cowa
A
di
A
Craig
be
ic
Craig
gr
pr
De M
is
pr
N
Good
m
w
J.
Good
m
pi
R

THIS PAGE BLANK (USPTO)

is at present unclear. It also remains to be established why it did not occur in cambial cells; when observed, the number of gold particles was within the range of background level. Although the molecular mechanism of plant hormones is not understood, ABA may act by affecting the synthesis of certain mRNAs and proteins. It has been proposed recently that ABA regulates at the translational level in germinating seeds (Rodriguez et al. 1985).

Many questions arise from these observations. Thus, more studies are now needed to explain the mechanism of deposition of ABA within cells at different times.

In conclusion, immunogold techniques using electron-microscopy are to be preferred when estimating very small concentrations of reaction products, as we have shown for cytosolic labelling with differentiated cells. Furthermore, no correction for section thickness must be made in such analyses whereas in thick sections there is the problem of accessibility of antibodies. These techniques appeared promising for a detailed in-situ localization of other plant regulators against which antibodies are produced.

We are indebted to Dr. Mapelli for the generous gift of PA and DPA. The assistance of Y. Habricot in the preparation of the manuscript is gratefully acknowledged. We are grateful to Drs. Ferrand and Vaudois for allowing us use of the electron microscope.

References

- Bendayan, M. (1984) Protein A-gold electron microscopic immunocytochemistry: methods, applications and limitations. *J. Electron Microsc. Technol.* 1, 243-270
- Carlemalm, E., Garavito, R.M., Villiger, W. (1982) Resin development for electron microscopy and an analysis of embedding at low temperature. *J. Microsc. (Oxford)* 126, 123-143
- Cowan, I.R., Raven, J.A., Hartung, W., Farquhar, G.D. (1982) A possible role for abscisic acid in coupling stomatal conductance and photosynthetic carbon metabolism in leaves. *Aust. J. Plant Physiol.* 9, 489-498
- Craig, S., Goodchild, D.J. (1982) Post-embedding immunolabelling. Some effects of tissue preparation on the antigenicity of plant proteins. *Eur. J. Cell Biol.* 28, 251-256
- Craig, S., Miller, C. (1984) LR white resin and improved on-grid immunogold detection of Vicilin, a pea seed storage protein. *Cell Biol. Int. Rep.* 8, 879-886
- De Mey, J. (1983) Colloidal gold probes in immunocytochemistry. In: *Immunocytochemistry: practical applications in pathology and biology*, vol. 1, pp. 82-113. Polak, J.M., Van Noorden, S., eds. Wright, Bristol
- Goodchild, D.J., Andersson, B., Anderson, J.M. (1985a) Immunocytochemical localization of polypeptides associated with the oxygen evolving system of photosynthesis. *Eur. J. Cell Biol.* 36, 294-298
- Goodchild, D.J., Anderson, J.M., Andersson, B. (1985b) Immunocytochemical localization of the cytochrome b/f complex of chloroplast thylakoid membranes. *Cell Biol. Int. Rep.* 9, 715-721
- Greenwood, J.S., Chrispeels, M.J. (1985) Immunocytochemical localization of phaseolin and phytohemagglutinin in the endoplasmic reticulum and Golgi complex of developing bean cotyledons. *Planta* 164, 295-302
- Harris, N., Croy, R.R.D. (1985) The major albumin protein from pea (*Pisum sativum* L.). Localization by immunocytochemistry. *Planta* 165, 522-526
- Hartung, W., Heilmann, B., Gimmmler, H. (1981) Do chloroplasts play a role in abscisic acid synthesis? *Plant Sci. Lett.* 22, 235-242
- Herman, E.M., Shannon, L.M. (1984) Immunocytochemical evidence for the involvement of Golgi apparatus in the deposition of seed lectin of *Bauhinia purpurea* (Leguminosae). *Protoplasma* 121, 163-170
- Leroux, B., Maldiney, R., Miginiac, E., Sossountzov, L., Sotta, B. (1985) Comparative quantitation of abscisic acid in plant extracts by gas-liquid chromatography and an enzyme-linked immunosorbent assay using the avidin-biotin system. *Planta* 166, 524-529
- Maldiney, R., Leroux, B., Sabbagh, I., Sotta, B., Sossountzov, L., Miginiac, E. (1986) Enzyme immunoassay to quantify three phytohormones: auxin, abscisic acid and zeatin-riboside, using a biotin/avidin system. *J. Immunol. Meth.*, in press
- Milborrow, B.V. (1979) Antitranspirants and regulation of abscisic acid content. *Aust. J. Plant Physiol.* 6, 249-254
- Newman, G.R., Jasani, B., Williams, E.D. (1983) A simple post-embedding system for the rapid demonstration of tissue antigens under the electron microscope. *Histochem. J.* 15, 543-555
- Rodriguez, D., Nicolas, G., Aldasoro, J.J., Hernandez-Nistal, J., Babiano, M.J., Matilla, A. (1985) Altered development of polysomal RNA activity in chick-pea (*Cicer arietinum* L.) embryonic axes. Effects of abscisic acid and temperature. *Planta* 164, 517-523
- Roth, J., Bendayan, M., Orci, L. (1978) Ultrastructural localization of intracellular antigens by the use of protein-A gold complex. *J. Histochem. Cytochem.* 26, 1074-1081
- Roth, J., Bendayan, M., Carlemalm, E., Villiger, W., Garavito, M. (1981) Enhancement of structural preservation and immunocytochemical staining in low temperature embedded pancreatic tissue. *J. Histochem. Cytochem.* 29, 663-671
- Shaw, P.J., Henwood, J.A. (1985) Immuno-gold localization of cytochrome f, light-harvesting complex, ATP synthase and ribulose 1,5-bisphosphate carboxylase/oxygenase. *Planta* 165, 333-339
- Sotta, B., Sossountzov, L., Maldiney, R., Sabbagh, I., Tachon, P., Miginiac, E. (1985) Abscisic acid localization by light microscopic immunohistochemistry in *Chenopodium polyspermum* L. *J. Histochem. Cytochem.* 33, 201-208
- Tomenius, K., Clapham, D., Oxelfelt, P. (1983) Localization by immunogold cytochemistry of viral antigens in sections of plant cells infected with red clover mottle virus. *J. Gen. Virol.* 645, 2669-2678
- Walton, D.C. (1980) Biochemistry and physiology of abscisic acid. *Annu. Rev. Plant Physiol.* 31, 453-489
- Yamamoto, N., Yasuda, K. (1977) Use of a water soluble carbodiimide as a fixing agent. *Acta Histochem. Cytochem.* 10, 14-37
- Zavala, M.E., Brandon, B.L. (1983) Localization of a phytohormone using immunocytochemistry. *J. Cell Biol.* 97, 1235-1239
- Zeevaert, J.A.D. (1977) Sites of abscisic acid synthesis and metabolism in *Ricinus communis*. *L. Plant Physiol.* 59, 788-791

Received 13 January; accepted 17 March 1986

THIS PAGE BLANK (USPTO)

The differential ultrastructural localization of immunoglobulin heavy and light chains in human haematopoietic cell lines

D. G. NEWELL, A. HANNAM-HARRIS, A. KARPAS* AND J. L. SMITH *Department of Haematology, Faculty of Medicine and Regional Immunology Unit, Southampton General Hospital, Southampton, and *Department of Haematological Medicine, Cambridge*

Received 10 June 1981; accepted for publication 19 August 1981

SUMMARY. The ultrastructural localization of immunoglobulin heavy and light chains has been investigated in nine haematopoietic cell lines, using a technique which involves the treatment of lightly prefixed cells with saponin to allow penetration of the antibody-peroxidase conjugate. The synthesis and secretion of immunoglobulin was also studied in these cell lines.

Immunoglobulin was found to be localized in the cisternae and on the membranes of the rough endoplasmic reticulum, perinuclear space and/or Golgi apparatus. In each case staining for heavy chains was weak or absent in the perinuclear space while staining for light chains was usually strong. Additionally in three cell lines immunoperoxidase staining indicated that heavy chains were absent from the Golgi apparatus despite the observed presence of light chains in the Golgi apparatus and the secretion of combined immunoglobulin into the supernatant. The results obtained suggest compartmentalization of the synthesis of light and heavy chains and indicate that the technique of immunoelectron microscopy may significantly contribute to an understanding of the mechanisms involved in immunoglobulin synthesis, intracellular transport and secretion.

Human cell lines derived from normal or neoplastic lymphoid cells have been classified into four groups, lymphoblastoid, lymphoma, lymphoid and myeloma on the basis of cell source, morphology and functional parameters (Nilsson & Ponten, 1975). Although lymphoid T cell lines have been described (Minowada *et al.* 1972), the majority of haematopoietic cell lines are B cell in origin and synthesize immunoglobulin (Finegold *et al.* 1968).

Recent evidence suggests that B cell lines from lymphoid tissue sources are heterogeneous

Correspondence: Dr D. G. Newell, Department of Haematology, Faculty of Medicine, Southampton General Hospital, Southampton SO9 4XY.

0007-1048/82/0300-0445\$02.00 © 1982 Blackwell Scientific Publications

THIS PAGE BLANK (USPTO)

in terms of size, morphology, function and expression of immunoglobulin and other surface markers (Nilsson & Ponten, 1975; Gordon *et al.* 1978). These B lymphocyte populations may represent cells at different stages in the maturation of the B cell from pre-B lymphocytes to immunoglobulin secreting plasma cells. Lymphoid cell lines therefore may be used as models of normal B cell function and differentiation especially with regard to immunoglobulin synthesis, expression and secretion.

Biochemical investigations on immunoglobulin synthesis have indicated that heavy and light chains are synthesized on separate membrane bound ribosomes (Uhr, 1970) and sequestered into the cisternae of the rough endoplasmic reticulum (RER) from which they are transported into the smooth membranes of the Golgi apparatus (Zagury *et al.* 1970). The combination of light chains, from a free pool in the cisternae, with heavy chains still attached to the ribosomes may or may not be a prerequisite for heavy chain release (Vassalli *et al.* 1971). The intracellular transport of immunoglobulin from the RER through the Golgi cisternae is accompanied by the stepwise addition of carbohydrate moieties which is completed just prior to, or at the time of, secretion into the supernatant (Choi *et al.* 1971; Melchers & Andersson, 1974).

Previous attempts to visualize the cellular processes involved in immunoglobulin synthesis, intracellular transport and secretion or surface expression have been severely restricted by inadequate morphological preservation, poor conjugate penetration or lack of immunoglobulin class specificity of the labelled antibodies used. The mechanisms involved are therefore still obscure, especially in those cells containing small amounts of immunoglobulin which are preferentially incorporated into the membrane rather than secreted.

During this investigation we have utilized a technique recently developed to observe the ultrastructural localization of immunoglobulin in nine human lymphoid cell lines (Newell *et al.* 1980) and correlated this information with the immunoglobulin biosynthesis determined by the incorporation of ^3H -leucine. The results obtained suggest that marked differences in the distribution of heavy and light chains exist in these cell lines. The relationships between these differences, the cell source and state of maturation is discussed.

MATERIALS AND METHODS

Cell lines. Cell lines Jijoye and EB2 were purchased from Flow Laboratories (Irvine, Scotland). Cell line U266BL was kindly provided by Dr K. Nilsson (Uppsala, Sweden). The origin and characteristics of cell lines Karpas 160, Karpas 119 and Karpas 129 have been described previously (Gordon *et al.* 1977). Cell lines Karpas 328 and Karpas 427 were established from a chronic myeloid leukaemia and acute myeloid leukaemia respectively and have been in culture for over 24 months. All cell lines were cultured in RPMI 1640 (Biocult-Gibco Ltd) with 10% fetal calf serum (Sera Labs Ltd) containing 50 units benzyl penicillin/ml and 10 mg/ml streptomycin and harvested 24 h after splitting whilst in the log phase of growth.

Immunoglobulin expression

Immunofluorescence. Cell suspensions were stained with fluorescein conjugated rabbit or

sheep anti-
preparati
by the di
Orthoplai

Electr
Ficoll/Tri
Sorenson
then trea
aliquots o
sheep anti
glutarald
prepared
the pero:
(1966).
buffer, pH
30 min a
in Spurr
TEM at 4

Immu
sized by
1-[4,5- ^3H]
(1977).
gel electr
classes v
was che
class ex
immunc

Immunoj

Table I s
immunc
always c
cence in
consider
was not

Synthesi

All cell l
of the t
superna

THIS PAGE BLANK (USPTO)

sheep antiserum to human immunoglobulin heavy and light chains. Cytocentrifuged cell preparations were fixed in methanol and washed in saline before staining with these antisera by the direct method. The fluorescein-labelled preparations were examined using a Leitz Orthoplan microscope fitted with a HB 200 mercury vapour Ploem illuminator.

Electron microscopic immunoperoxidase. Non-viable cells were firstly removed on a Ficoll/Triosil gradient. Viable cells were then lightly fixed in 0.1% glutaraldehyde in 0.1 M Sorenson's phosphate buffer, pH 7.2, for 15 min at room temperature. The fixed cells were then treated with 1% saponin (Sigma Chemical Co. Ltd) for 10 min at 55°C. After washing, aliquots of cells (1×10^7 cells) were incubated for 1 h with 100 μ l of 10 mg/ml affinity purified sheep anti-human immunoglobulin conjugated to horse radish peroxidase by the two-step glutaraldehyde method (Avrameas & Ternyck, 1971). (Control conjugates were also prepared using normal sheep IgG.) The cells were washed to remove unbound conjugate and the peroxidase developed using the diaminobenzidine substrate of Graham & Karnovsky (1966). The cells were then post-fixed in 1% glutaraldehyde, 0.1 M Sorenson's phosphate buffer, pH 7.2, 15 min at room temperature, followed by 1% osmic acid in Palade's buffer, for 30 min at room temperature. The stained cells were embedded in 2% agar before embedding in Spurr resin. Ultra-thin sections were cut using glass knives and viewed on a Phillips 201 TEM at 40 kV.

Immunoglobulin biosynthetic studies. The amount and class of immunoglobulin synthesized by the cells over an 18 h culture period was determined by the incorporation of L-[4,5- 3 H]leucine followed by immunoprecipitation according to the method of Gordon *et al* (1977). The immunoglobulin synthesized was also characterized using SDS-polyacrylamide gel electrophoresis. Specific antisera to labelled immunoglobulin light chain and heavy chain classes were raised in sheep and prepared by affinity purification and absorption. Specificity was checked by the ability to precipitate myeloma proteins of a single heavy or light chain class exclusively. Anti-free light chain antibodies retained no activity for combined immunoglobulin.

RESULTS

Immunofluorescence data

Table I shows that all cell lines except Raji stained for surface immunoglobulin using direct immunofluorescence with FITC labelled antisera. When present the intracellular Ig class always correlated with that of the surface. Intracellular Ig was detected by immunofluorescence in six out of nine cell lines and the intensity and proportion of staining varied considerably between cell lines. Cell line Karpas 427 expressed both κ and λ light chains and was not considered to be monoclonal.

Synthesis data

All cell lines were studied for Ig synthesis (Table II). Ig production is expressed as a percentage of the total protein synthesized in the lysate and supernatant. The Ig secreted into the supernatant is expressed as a percentage of the total Ig produced. In all cases the Ig class

THIS PAGE BLANK (USPTO)

Table I. Lymphoid cell lines: origin and Ig expression

Cell line	Original disorder	EBV	Ig expression*	
			SIg	IcIg
Raji	BL	+	0	0
Jijoye	BL	+	0	0
EB2	BL	+	100 G λ	30 G λ
Karpas 129	AML	+	100 G κ	0
Karpas 119	AML	+	60 M λ	60 M λ
Karpas 328	CML	+	47 A κ	30 A κ
Karpas 160	ALL	+	100 MD λ	20 M λ
Karpas 427	AML	+	41 M κ (λ)	65 M κ (λ)
U266BL	Myeloma	-	16 E λ	30 E λ

SIg = surface immunoglobulin; IcIg = intracellular immunoglobulin; BL = Burkitt's lymphoma; AML = acute myeloblastic leukaemia; CML = chronic myelocytic leukaemia; ALL = acute lymphoblastic leukaemia.

*% positive by immunofluorescence.

Table II. Immunoglobulin synthesis by cell lines

Cell line	Ig synthesized (% total protein)	Ig secreted (% total Ig)	LCf sup (% Ig sup)	LCf Lys (% Ig Lys)	Ig class
Raji	1.0	0	0	0	0
Jijoye	1.4	45	70	90	κ
EB2	1.4	65	10	28	G λ
Karpas 129	0.6	45	71	ND	G κ
Karpas 119	2.0	41	41	24	M λ (D)
Karpas 328	2.6	55	20	0	A κ
Karpas 160	1.0	15	0	30	M λ (D)
Karpas 427	0.5	54	22 λ 11 κ	ND	G (M) λ (κ)
U266BL	10.6	57	48	ND	E λ

LCf = Ig precipitated with anti free light chain antibodies; Ig = Ig precipitated with anti Fab γ antibodies; Lys = lysate; sup = supernatant.

produced was consistent with the fluorescence data. In the cases of Raji and U266BL the lysates and supernatants were analysed by polyacrylamide gel electrophoresis and these gel analyses confirmed the synthesis studies. Free light chain production was detected using anti-free light chain antibodies and was expressed as the percentage of the total material precipitated by the anti Fab γ in the supernatant or lysate.

THIS PAGE BLANK (USPTO)

Immunoperoxidase localization of Ig

The class of Ig located intracellularly in these cell lines correlated with the immunofluorescent and/or biosynthesis data except in Raji where κ light chains were detected by immunoperoxidase but not biosynthetically and in Karpas 129 where no G heavy chains were detected (Table III) by immunoperoxidase.

Table III. Percentage of cells staining for Ig and the intensity of staining in cell lines

	Heavy chain					Light chain				
	Class	% +ve	PNS	RER	Golgi	Class	%	PNS	RER	Golgi
Raji	M	50	±	+	—	κ	30	±	±	—
Jijoye	O	0	—	—	—	κ	50	+	+	+
EB2	G	50	±	+	+	λ	90	+	+	+
Karpas 129	O	0	—	—	—	κ	75	±	+	±
Karpas 119	M	40	±	+	—	λ	70	+	+	+
Karpas 328	A	29	+	+	—	κ	55	+	+	+
Karpas 160	M	45	—	+	+	λ	95	+	+	+
Karpas 427	M	8	—	±	—	κ(λ)	63(31)	+	+	+
U266BL	E	ND				λ	55	+	+	+

Staining occurred in the cisternae and/or the membranes of the RER and of the perinuclear space and in the area of the Golgi apparatus. Staining localization and intensity varied considerably both between cell lines and within cell lines. In cell lines Karpas 129, Karpas 119, Karpas 160 and Karpas 328 at least 50 cells were observed and the staining pattern and intensity noted (Table IV). No attempt was made to correlate localization with cell cycle but mitotic figures showed some staining in RER strands. Staining of surface immunoglobulin (sIg) is not observed using this technique probably because sIg is wholly or partially removed by the saponin treatment (Newell *et al.* 1980).

Perinuclear space

Staining for heavy chains in the perinuclear space was, in all cases except Karpas 328, weak or absent. Conversely, the staining for light chains was usually strong, sometimes being patchy or localized in a small area (Figs 1 and 2). The staining in the PNS frequently showed the contiguous nature of the PNS and RER cisternae.

Rough endoplasmic reticulum

Usually both heavy and light chains were demonstrable in the RER. The patterns of RER distribution were strikingly different in some cells when stained, i.e. Karpas 328 consisted of

THIS PAGE BLANK (USPTO)

Table IV. Percentage distribution and intensity of staining for immunoglobulin

Cell line	Ig class	Staining intensity								
		PNS			Golgi			RER		
		+	±	-	+	±	-	+	±	-
Karpas 129	κ	46	9	43	25	20	54	64	29	6
Karpas 119	λ	74	19	8	71	19	9	82	14	2
	μ	20	24	56	5	0	95	38	61	4
Karpas 160	λ	57	26	16	54	11	34	88	11	0
	μ	0	0	100	14	23	63	72	27	0
Karpas 328	κ	48	48	3	11	33	55	52	41	5
	A	76	15	8	0	0	100	78	16	5

large rings of stain apparently enclosing areas of cytoplasm (Fig 3a). EB2 had long strands forming an interconnecting network (Fig 2a) and U266BL showed small vesicles of stain throughout the cytoplasm (Fig 4). Frequently the stained strands of RER ran up to and appeared to merge with the cell membrane and in U266BL the stain vesicles were seen around the periphery of the cell occasionally located at one end of the cell only (Fig 4).

In those cell lines showing morphological evidence for surface projections, positive staining could be associated with these projections in one line only (U266BL). In Karpas 160 surface protuberances were observed containing peroxidase activity associated with strands of RER (Fig 5). In some cell lines not all the RER was stained (Karpas 328, Fig 3). This did not appear to be associated with any particular location within the cell and was therefore not considered to be the result of inadequate conjugate penetration.

In Raji (Fig 6) and Karpas 160 staining appeared to be preferentially located along the membranes of the RER rather than intracisternal, producing a tram-line effect. Similar staining was seen along the PNS membranes in these cell lines. Karpas 328 and U266BL conversely had heavy staining throughout the RER cisternae while the rest of the cells showed varying degrees of cisternal and membranous staining.

Golgi apparatus

Staining was not always associated with the Golgi apparatus, even in cell lines which were actively secreting immunoglobulin. In three cell lines, Karpas 119, Karpas 427 and Karpas 328 (Fig 3), an asynchrony between the light chain and heavy chain localization in the Golgi was observed. In these cell lines very weak or negligible staining was observed in the Golgi apparatus with anti-heavy chain conjugates, whilst the anti-light chain conjugates produced relatively strong staining. When present, the staining in the Golgi area was frequently diffuse

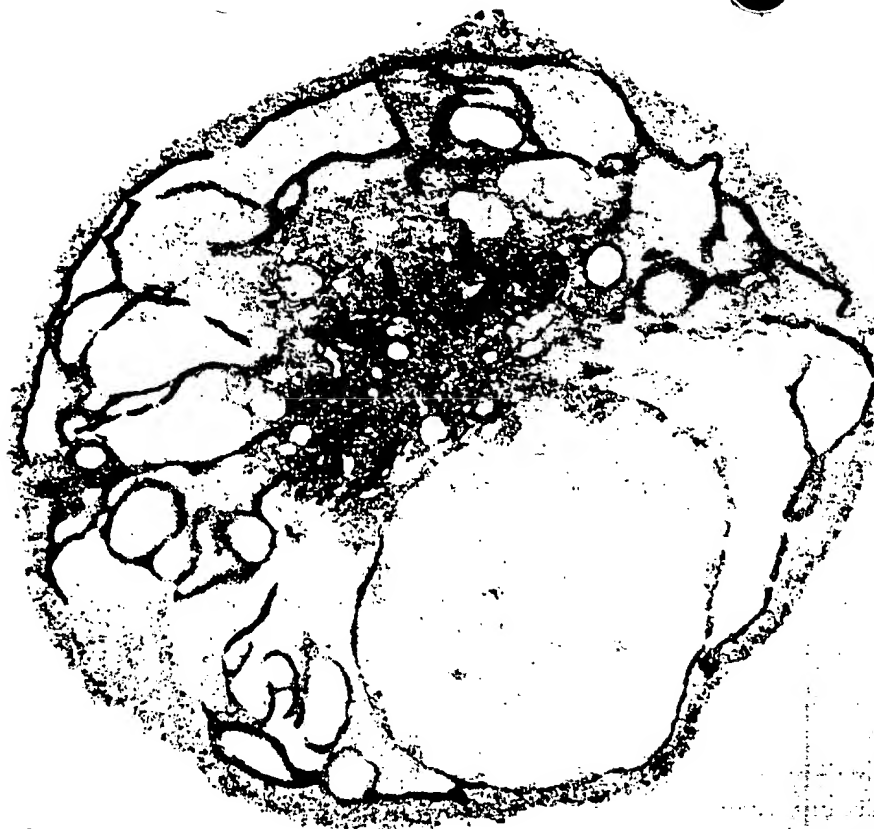
Fig 1. Ce
absence
chains. (

THIS PAGE BLANK (USPTO)



Fig 1. Cells from cell line Karpas 160 labelled (a) for M heavy chain and (b) λ light chain. Note the absence of staining for immunoglobulin M in the perinuclear space compared with the presence of λ chains. (a) $\times 12\ 000$; (b) $\times 11\ 500$.

THIS PAGE BLANK (USPTO)



a



b

Fig 2. Cells from cell line EB2 labelled (a) for G heavy chain and (b) for λ light chain. (a) $\times 10\ 000$; (b) $\times 10\ 500$.

Fig 3. (a) absence of label (a) $\times 10\ 000$

THIS PAGE BLANK (USPTO)

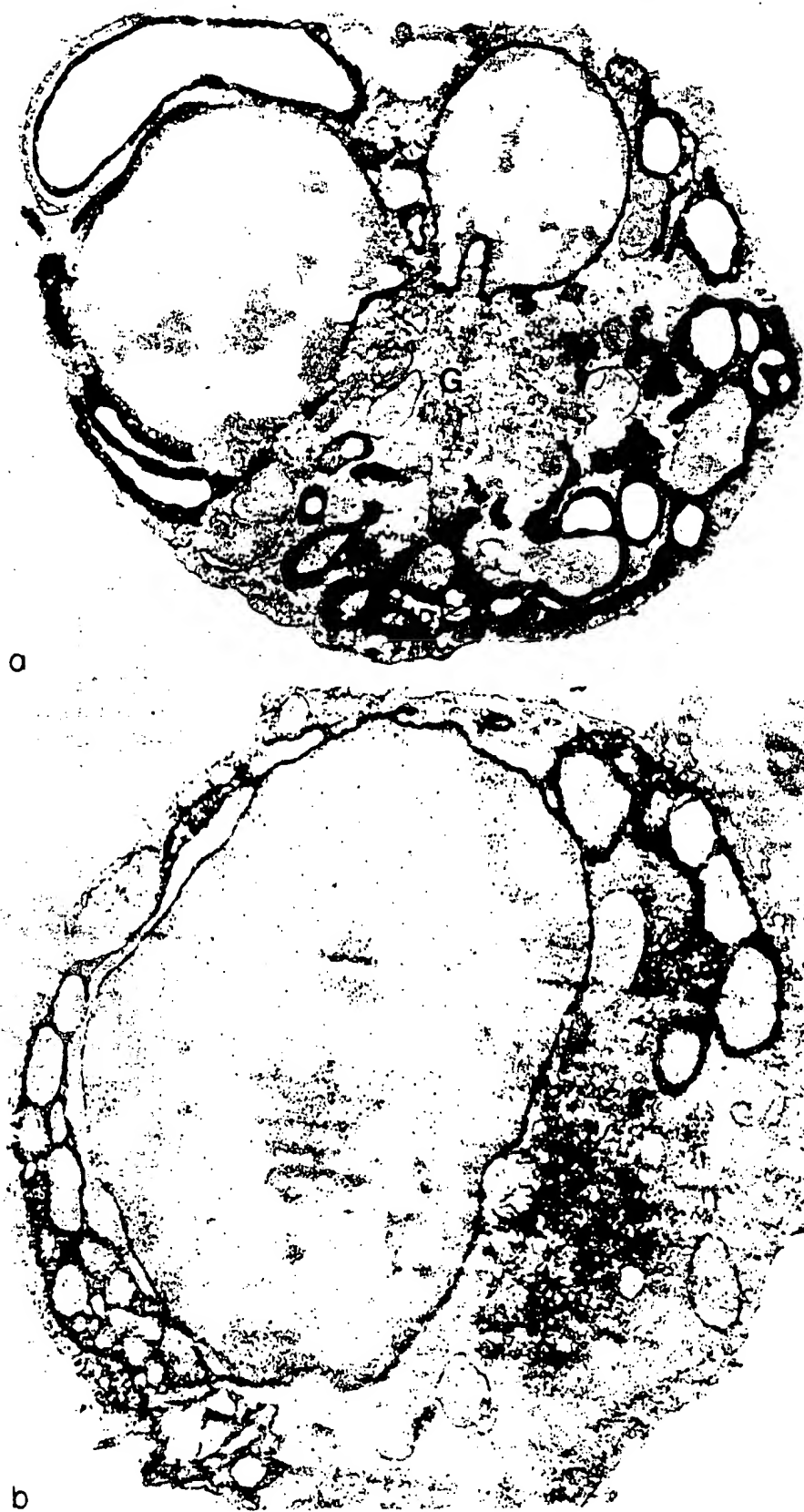


Fig 3. Cells from cell line Karpas 328 labelled (a) for A heavy chain and (b) λ light chain. Note the absence of immunoglobulin A in the Golgi apparatus (G) while the staining for λ chains is quite strong. (a) $\times 12\,600$; (b) $14\,000$.

THIS PAGE BLANK (USPTO)



Fig 4. Cell line U266BL stained for λ light chains showing the presence of small vesicles of RER containing stained material in the cytoplasm and similar vesicles apparently being secreted from the cell (↑). $\times 9500$.

rather than located specifically in the vacuoles. In some cases the stain was obviously localized on the membranes, whilst in others it filled the vacuoles.

DISCUSSION

The localization of Ig in nine cell lines has been investigated. In each case the monoclonality found by TEM-immunoperoxidase corresponded with that obtained by immunofluorescence and ^3H -leucine incorporation. In one case (Raji) κ light chains were detected ultrastructurally but not by the biosynthetic technique. This staining was monoclonal and restricted to the cisternae of the PNS and RER. The expression of κ light chains by Raji cells has been previously indicated (Takahashi *et al*, 1969; Klein *et al*, 1977) although some Raji cell line cultures express λ light chains (Preud'homme *et al*, 1978). The presence of κ light chains detected by the TEM-immunoperoxidase technique and not the ^3H -leucine incorporation may be explained by an increased sensitivity and/or a slow rate of biosynthesis.

In cell line Karpas 129 the γ heavy chains detected biosynthetically were not localized by TEM-immunoperoxidase. It is possible that the anti γ conjugate was not as efficient as the other conjugates although it successfully detected γ heavy chains in the EB2 cell line.

Fig 5. Cell
containin

In a
membr
consiste
previou
synthes
immun
reporte
found
conside
Occasic
Golgi a
to post
The
cell lin
analysi
related
Hodge

THIS PAGE BLANK (USPTO)

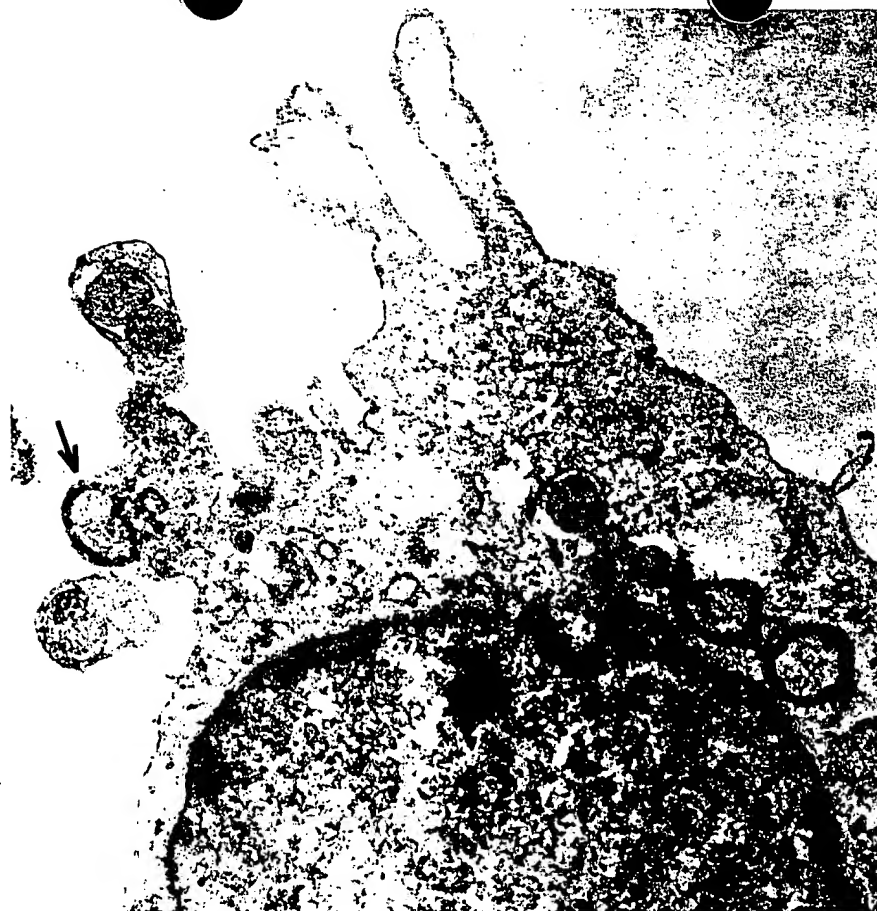


Fig 5. Cell line Karpas 160 stained for λ light chains showing surface protuberances with strands of RER containing stained material (\uparrow). $\times 18\ 000$.

In all the cell lines studied the intracellular Ig was restricted to the cisternae and membranes of the PNS and RER and the area of the Golgi apparatus. This distribution is consistent with the synthesis of immunoglobulin on membrane bound ribosomes which has previously been demonstrated by Uhr (1970). However, recent reports suggest that some synthesis on free polyribosomes is possible (Omori *et al.*, 1977) and the presence of immunoperoxidase staining for immunoglobulin on cytoplasmic ribosomes has been reported in several cell lines (Raji and T51) (Preud'homme *et al.*, 1978). Diffuse staining when found during our investigations was invariably non-monoclonal and was therefore considered to be artefactual and probably due to cell degeneration (Tsunoda *et al.*, 1978). Occasionally diffuse staining appeared in cytoplasmic areas around heavily stained RER or Golgi apparatus and this was considered to be a result of diffusion of the reaction product prior to post fixation.

The considerable variation observed in the intensity and patterns of staining within the cell lines may, in part, be due to cell cycle related changes in Ig synthesis. Biochemical analysis of synchronous lymphoid cells indicate that both Ig synthesis and secretion are related to the cell cycle and peak in late G1 and early S phase (Buell & Fahey, 1969; Lerner & Hodge, 1971) with a possible second peak in G2 phase (Watanabe *et al.*, 1972). Such

THIS PAGE BLANK (USPTO)



Fig 6. Cells from the cell line Raji stained (a) for M heavy chains and (b) for λ light chains. Note the localization of strain along the membranes of the RER rather than intracisternal. (a) $\times 10\,000$; (b) $\times 9800$.

changes of
the cell.
M phase,
mitosis.

One of
between h
in the PN
in prelim
peroxidas
termed pr
comprise
movement
membran
preferent
suggests
sequester
This may
(Gordon
mitogen
secretion
perinucle

Heav
molecula
that such
in the re

Differ
three ce
superna
transport
smooth
carbohy
al, 1971
detected
our stud
be close
have be
such as
or extre

No c
was obt
the plas
In the c
released



THIS PAGE BLANK (USPTO)

changes could easily be reflected in differences in Ig distribution and intensity within the cell. Although biochemical analysis has shown that Ig synthesis is minimal in M phase, RER cisternae containing Ig stained with peroxidase were found in cells in active mitosis.

One of the most striking results of this study is the apparent differences in localization between heavy and light chains. In most cases the heavy chain was absent, or much reduced, in the PNS while the staining for light chain was strong. Similar patterns have been observed in preliminary studies on chronic lymphocytic leukaemic cells (CLL). In studies using peroxidase as the immunogen, marker cells showing antibody preferentially in the PNS were termed primitive cells (Leduc *et al.*, 1968). Anti-peroxidase antibody in this case presumably comprised complete immunoglobulin not light chain only, but it does further suggest that movement of antibody production from the perinuclear cisternae to the cytoplasmic membrane bound ribosomes accompanies lymphocyte differentiation into plasma cells. The preferential localization of light chain in the PNS, in most of the cell lines investigated, suggests that light chain synthesis occurs primarily in the PNS and that free light chain, sequestered in the cisternae, is transported to the sites of heavy chain synthesis in the RER. This may well be related to the presence of excess free light chains in immature lymphoid cells (Gordon *et al.*, 1978; Gordon & Smith, 1978). Furthermore, CLL cells stimulated by pokeweed mitogen undergo a change from light chain excess to balanced heavy and light chain secretion, which is accompanied by a concomitant observation of heavy chains in the perinuclear space and Golgi apparatus (Newell, unpublished observation).

Heavy and light chains have been shown to be synthesized on polyribosomes of different molecular weights (i.e. on separate ribosomes) (Uhr, 1970) and our observations indicate that such synthesis is also compartmentalized. Such compartmentalization may be involved in the regulation of synthesis and prevention of premature polymerization.

Differences in heavy and light chain distribution in the Golgi apparatus were also noted in three cell lines, which was inconsistent with the secretion of combined antibody into the supernatant by these cell lines. Biochemical and autoradiographic analysis of Ig intracellular transport and secretion has indicated that the passage of newly formed Ig through the smooth membranes of the Golgi apparatus is accompanied by the addition of terminal carbohydrate moieties and is probably a prerequisite for secretion (Zagury *et al.*, 1970; Choi *et al.*, 1971). In some early studies localization of immunoglobulin in the Golgi could not be detected (Suzuki *et al.*, 1970) but this may have been due to poor conjugate penetration. In our studies the presence of light chain, but not heavy chain, in the Golgi apparatus appears to be closely associated with the secretion of free light chains. Combined immunoglobulin may have been undetectable in the Golgi apparatus either because of secretion via another route such as clasmotaxis (Thiery, 1960) or diffusion to the plasma membrane (Leduc *et al.*, 1968) or extremely rapid passage through the smooth endoplasmic reticulum.

No conclusive evidence for the mechanism of immunoglobulin secretion in these cell lines was obtained. In most cases runs of RER containing immunoglobulin appeared to merge with the plasma membrane and occasionally these strands were found in surface protuberances. In the case of the myeloma cell line U266BL, small vesicles of stained material appeared to be released from the cell surface. It is therefore most likely that several mechanisms for

THIS PAGE BLANK (USPTO)

immunoglobulin release exist within the same cell dependent probably on the immunoglobulin involved and the state of maturation of the cell.

The localization of immunoglobulin on the membranes and/or in the cisternae of the RER suggests that some association between intracellular localization of immunoglobulin and surface expression or secretion may exist. Preliminary investigations on CLL cells expressing surface immunoglobulin without secretion, indicate that the intracellular location of immunoglobulin on membranes is concomitant with surface membrane expression. It is therefore likely that the integration of immunoglobulin into membranes takes place at or about the time of synthesis.

ACKNOWLEDGMENTS

Financial support for this project was provided by the Cancer Research Campaign and the Leukaemia Research Fund of the Wessex Regional Health Authority.

REFERENCES

- AVRAMEAS, T. & TERNYCK, S. (1971) Peroxidase labelled antibody and Fab conjugates with enhanced intracellular penetration. *Immunochimistry*, **8**, 1175-1179.
- BUELL, D.N. & FAHEY, S.L. (1969) Limited periods of gene expression in immunoglobulin synthesizing cells. *Science*, **164**, 1524-1525.
- CHOI, Y.S., KNOOF, P.M. & LENNOX, E.S. (1971) Intracellular transport and secretion of an immunoglobulin light chain. *Biochemistry*, **10**, 668-679.
- FINEGOLD, I., FAHEY, J.L. & DUTCHER, T.F. (1968) Immunofluorescent studies of immunoglobulins in human lymphoid cells in continuous culture. *Journal of Immunology*, **101**, 366-373.
- GORDON, J., HOUGH, D., KARPAS, A. & SMITH, J.L. (1977) Immunoglobulin expression and synthesis by human haemic cell lines. *Immunology*, **32**, 559-565.
- GORDON, J., HOWLETT, T.R. & SMITH, J.L. (1978) Free light chain synthesis by neoplastic cells in chronic lymphocytic leukaemia and non-Hodgkin's lymphoma. *Immunology*, **32**, 397-404.
- GORDON, J. & SMITH, J.L. (1978) Free immunoglobulin light chain synthesis by neoplastic cells in leukaemic reticuloendotheliosis. *Clinical and Experimental Immunology*, **31**, 244.
- GRAHAM, R.C. & KARNOVSKY, M.J. (1966) The early stages of absorption of injected horseradish peroxidase in the proximal tubules of mouse kidney: ultrastructural cytochemistry by a new technique. *Journal of Histochemistry and Cytochemistry*, **14**, 291-302.
- KLEIN, G., TARASAKI, P., BILLING, R., HONIG, R., JONDAL, M., ROSEN, A., ZEULHEN, J. & CLEMENTS, G. (1977) Somatic cell hybrids between human lymphoma lines. III. Surface markers. *International Journal of Cancer*, **19**, 66-76.
- LEDUC, F., AVRAMEAS, S. & BOUTEILLE, M. (1968) Ultrastructural localization of intracellular immunoglobulins in plasma cells and lymphoblasts by enzyme labelled antibodies. *Journal of Experimental Medicine*, **127**, 109-118.
- LERNER, R.S. & HODGE, L.D. (1971) Gene expression in synchronised lymphocytes: studies on the control of synthesis of immunoglobulin polypeptides. *Journal of Cell Physiology*, **77**, 265-276.
- MELCHERS, F. & ANDERSSON, J. (1974) Immunoglobulin production in B lymphocytes: synthesis of the membrane-bound receptor and the secreted serum glycoprotein immunoglobulin M. *Biochemical Society Symposium*, **40**, 73-85.
- MINOWADA, J., OHNUMA, T. & MOORE, G.E. (1972) Rosette forming human lymphoid cell lines. 1. Establishment and evidence for origin of thymus-derived lymphocytes. *Journal of the National Cancer Institute*, **49**, 891-895.
- NEWELL, D.G., BOHANE, C., PAYNE, J. & SMITH, J.L. (1980) The intracellular localization of im-

munoglob
haematop
electron
Methods,
NILSSON, K.
biological
topoietic
Cancer, 1
OMORI, K.,
Immunog
somes se
City Medi
PREUD'HOMI
FELLOUS.
B cell cha
entiation-
Symposiu
Rosenfel
land, An
SUZUKI, I.,
MOTO, T
immuno
tissue cu
detected
Immunol
TAKAHASH
MAN, D
product

THIS PAGE BLANK (USPTO)

- munoglobulin in human lymphoid cells and haematopoietic cell lines by immunoperoxidase electron microscopy. *Journal of Immunological Methods*, 37, 275-286.
- NILSSON, K. & PONTEN, J. (1975) Classification and biological nature of established human haematopoietic cell lines. *International Journal of Cancer*, 15, 321-341.
- OMORI, K., WAKE, K. & MORISAWA, I. (1977) Immunoglobulin synthesis in free polyribosomes separated from a myeloma cell. *Osaka City Medical Journal*, 24, 33-50.
- PREUD'HOMME, J.L., GOURDIN, M.F., REYES, F. & FELLOUS, M. (1978) Lymphoid cell lines with pre B cell characteristics. *Human Lymphocyte Differentiation—Its Application to Cancer*, INSERM Symposium No. 8 (ed. by B. Serou and C. Rosenfeld), pp. 345-351. Elsevier/North Holland, Amsterdam.
- SUZUKI, I., TAKAHASHI, M., KAMEI, H. & YAMAMOTO, T. (1970) Intracellular distribution of immunoglobulin heavy and light chains within tissue cultured cells of human lymphoid origin detected by electron microscopy. *Journal of Immunology*, 104, 907-917.
- TAKAHASHI, M., YAGI, Y., MOORE, G.E. & PRESSMAN, D. (1969) Pattern of immunoglobulin production in individual cells of human haematopoietic origin in established culture. *Journal of Immunology*, 102, 1274-1283.
- THIERY, J.P. (1960). In: *Ciba Foundation Symposium on Cellular Aspects of Immunology* (ed. by G. E. W. Wolstenholm and M. O'Conner), p. 59. Churchill, London.
- TSUNODA, R., TERASHIMA, K., TAKAHASHI, K. & KOJIMA, M. (1978) An ultrastructural study with enzyme labelled antibody technique on immunoglobulin containing cells in human tonsil especially in germinal centres. *Acta Pathologica Japonica*, 28, 53-75.
- UHR, J.W. (1970) Intracellular events underlying synthesis and secretion of immunoglobulin. *Cellular Immunology*, 1, 228-244.
- VASSALLI, P., LISOWSKA-BERNSTEIN, B. & LAMM, M.E. (1971) Cell free synthesis of rat immunoglobulin. III. Analysis of the cell free made chains and of their mode of assembly. *Journal of Molecular Biology*, 56, 1-19.
- WATANABE, S., YAGI, Y., TANIGAKI, N. & PRESSMAN, D. (1972) Immunoglobulin synthesis in synchronised human lymphoid cell culture. *Federation Proceedings*, 31, 795.
- ZAGURY, D., UHR, J., JAMIESON, J. & PALADE, G. (1970) Immunoglobulin synthesis and secretion. *Journal of Cell Biology*, 46, 52-63.

THIS PAGE BLANK (USPTO)

BRITISH JOURNAL OF

ISSN 0007-1048

HAEMATOLOGY

VOL 50, NO 3, MARCH 1982



EDITED BY I. CHANARIN

Blackwell Scientific Publications

LANE MEDICAL LIBRARY

MAR 25 1982

STANFORD UNIVERSITY

THIS PAGE BLANK (USPTO)

A V region determinant (idiotope) expressed at high frequency in B lymphocytes is encoded by a large set of antibody structural genes

Renate Dildrop*, Jutta Bovens, Miriam Siekevitz,
Konrad Beyreuther and Klaus Rajewsky

Institute for Genetics, University of Cologne, Weyertal 121, D-5000 Köln
41, FRG

*To whom reprint requests should be sent

Communicated by K. Rajewsky

Idiotope Ac38, a V region determinant of the $\lambda 1$ chain-bearing, germ line encoded antibody B1-8, is expressed at high frequency ($\sim 1/40$) in $\lambda 1$ chain-bearing B cells. Here, we describe the isolation of lambda-positive hybridomas from C57BL/6 mice which had been immunized with antibody Ac38, the antibody recognizing idiotope Ac38. In Northern blot analysis, mRNA isolated from 10 such hybridomas hybridizes with a cDNA probe from the V_H gene expressed in the cell line B1-8. Amino acid sequence analysis of the V_H regions of four of the hybridoma proteins reveals that they are all derived from related, though distinct, germ line V_H genes. In one case the sequence data suggest that extensive somatic mutation has taken place. Only one of the four sequences derives from the same V_H gene that is expressed in the cell line B1-8. Together with earlier evidence, the present data demonstrate that the Ac38 idiotope is a marker for at least five V_H and three D region genes in the C57BL/6 germ line. This explains the high frequency at which this idiotope is expressed in the B cell population. In addition, our sequence determinations identify two V_H genes in the C57BL/6 strain which are closely related (and possibly allelic) to two known BALB/c V_H genes. One of these genes is the gene expressed in the BALB/c myeloma MOPC 104E.

Key words: V_H gene/hybridomas/B lymphocytes/idiotope expression

Introduction

Antibodies can interact with each other through variable (V) region determinants (idiotopes). It is clear from previous work that genetic polymorphism exists at the level of idiotopes and reflects polymorphism of antibody structural genes in the germ line. It has also been shown that idiotopes often depend upon the contribution of both the V_L and the V_H region, and that distinct portions of a given chain [e.g., the second and third complementary determining region (CDR) in the case of the heavy chain] may together be involved in the construction of a particular idiotope (reviewed by Rajewsky and Takemori, 1983). The genetic basis of idiotope expression is thus complex, and the frequency at which an idiotope is expressed in the B cell population may depend to a large extent upon the multiplicity of genetic elements (V, D and J regions) that can participate in the control of V regions expressing that idiotope.

In previous work (Reth *et al.*, 1979; Rajewsky *et al.*, 1981) a set of distinct determinants (idiotopes) was defined by monoclonal anti-idiotope antibodies on the V region of

antibody B1-8. Antibody B1-8 is of C57BL/6 origin and expresses a germ line-encoded V region (Bothwell *et al.*, 1981, 1982) with specificity for the hapten (4-hydroxy-3-nitrophenyl)acetyl (NP). The expression of some of the B1-8 idiotopes which are regularly detected on a fraction of the anti-NP antibodies produced in C57BL/6 mice upon immunization with NP was analyzed at the level of the B cell population by limiting dilution analysis (Takemori *et al.*, 1982; Nishikawa *et al.*, 1983). One of the idiotopes, namely idiotope Ac38, was found in the B cell population at a frequency of $\sim 1 \times 10^{-3}$. Since idiotope Ac38 is expressed almost exclusively on antibodies bearing $\lambda 1$ chains (Tesch *et al.*, 1983; C. Kappen, M. Reth and K. Rajewsky, in preparation) and B cells expressing $\lambda 1$ chains represent only a few percent of all B cells (Takemori and Rajewsky, 1981) this implies that approximately every 40th $\lambda 1$ -bearing B cell expresses idiotope Ac38. Only one out of 10 of these cells synthesizes an antibody with specificity for the NP hapten (Takemori *et al.*, 1982; Nishikawa *et al.*, 1983; Tesch *et al.*, 1983). The high frequency of Ac38 idiotope expression is also found in B lymphocytes maturing from bone marrow pre-B cells *in vitro* (Nishikawa *et al.*, 1983). Therefore, the high frequency of Ac38 expression is probably not due to selection of the corresponding cells after their generation in the bone marrow. Are certain V, D and J segments preferentially recombined or is the Ac38 idiotope expressed by V regions encoded by a variety of V, D and J segments?

Most $\lambda 1$ chains of murine antibodies express the same germ line-encoded variable region (Weigert and Riblet, 1976). This variable region is compatible with, but not sufficient for, the expression of idiotope Ac38, since the idiotope is not expressed by a variety of antibodies carrying germ line-encoded $\lambda 1$ chains (see Rajewsky and Takemori, 1983). The high frequency of cells bearing antibodies with both $\lambda 1$ chains and the Ac38 idiotope can therefore be understood by analyzing only the heavy chain variable regions of such antibodies.

Here we analyze the structural basis of Ac38 expression by sequencing the VDJ regions of four hybridoma proteins obtained from an immunization of C57BL/6 mice with the anti-idiotope antibody which recognizes idiotope Ac38. Such an immunization leads to the activation of Ac38-bearing B cells irrespective of their antigen binding specificity (Takemori *et al.*, 1982; Tesch *et al.*, 1983). The resulting Ac38-positive hybridomas are thus in the majority unreactive with the NP hapten. We have selected for the present study three hybridomas of the latter type and one which expresses NP binding specificity. Our results show that all four proteins express different V_H and D regions and that the V_H region of the NP-binding antibody corresponds to the germ line V_H gene which has previously been found to control a major fraction of the primary anti-NP response in C57BL/6 mice (Bothwell *et al.*, 1981; see also Rajewsky and Takemori, 1983).

THIS PAGE BLANK (USPTO)

Table I. Lambda I chain-bearing hybridomas induced by anti-idiotope Ac38

Hybridoma line	Fusion no.	Heavy chain	Binding to anti-idiotope antibody ^a							Binding to NIP ^a
			Ac38	Ac146	As79	A6-24	A25-9	A39-40	A31-90	
251.5	5	γ 1	+	- ^a	-	-	-	-	-	-
8.4	6	γ 1	+	-	-	+	± ^a	-	-	+
103.10	7a	γ 1	+	-	-	±	-	-	-	-
141.3	7a	γ 1	±	-	-	-	-	-	-	-
205.12	7b	γ 2b	±	-	-	-	-	-	-	-
214.8	7b	γ 1	±	-	-	-	-	-	-	-
225.4	7b	γ 1	+	-	-	-	-	-	-	-
260.2	7b	γ 1	+	-	-	-	-	-	-	-
303.20	7b	μ	+	-	-	±	-	-	-	+

^aBinding was determined on plastic plates coated with one of the various anti-idiotope antibodies or NIP₁₄-BSA.

+: indicates that plateau binding was $\geq 75\%$ of that observed with the control antibody B1-8 (in the case of antibody 8.4 binding to the anti-idiotope antibody Ac38 the value was 50%).

±: plateau binding 20–30% of that observed with antibody B1-8.

-: maximal binding < 10% of the B1-8 control.

Results

Production of Ac38-positive hybridoma cell lines

Spleen cells from C57BL/6 mice which had been immunized with Ac38-KLH, as described in Materials and methods, were fused to X63.Ag8.653 myeloma cells. After dilution and growth in microtiter wells, supernatants were screened in a plastic plate radioimmunoassay for λ 1-positive antibodies which could bind to antibody Ac38. We selected λ 1-positive antibodies because we wanted to examine idiotope Ac38-bearing antibodies as opposed to antibody Ac38-recognizing antibodies (anti-anti-idiotope antibody). Since >95% of all murine antibodies bear κ light chains (Takemori and Raskovsky, 1981), we would expect most anti-anti-idiotopic antibodies to be κ -positive. In contrast, we would also expect that λ 1-positive antibodies would be predominantly idiotope bearing. λ -Positive hybridomas were isolated from four separate fusions and cloned. Ten of these cell lines were randomly selected for further analysis.

In Table I, we list the hybridoma lines and give some serological properties of their antibody products. Most of the antibodies belonged to the IgG1 class, but an IgM and an IgG2b antibody were also present. Only two of the antibodies bound to the hapten (4-hydroxy-3-nitro-5-iodo-phenyl)acetyl (NIP). (Antibody B1-8 binds NIP with an affinity higher than that for the NP hapten against which it had been raised.) All antibodies bound to insolubilized antibody Ac38 (the property for which they had been selected), and in most instances this binding reached plateau levels close to the level reached by antibody B1-8 against which antibody Ac38 had originally been raised. In three of the cases maximal binding was considerably lower, indicating a lower affinity of these antibodies to antibody Ac38. The antibodies in Table I were also screened for the expression of six other B1-8 idiotopes. Four of these idiotopes were not found on any of the 10 antibodies. One idiotope (A25-9) was weakly expressed on one of the proteins and another one (A6-24) on four proteins, although maximal binding to the corresponding anti-idiotope was low in three out of four cases.

Hybridization of poly(A)⁺ RNA to B1-8 V_H cDNA

We wished to examine the genetic and structural basis for the similarities and differences between the proteins described in Table I. Since λ 1 expression is essentially associated with a single V_L and J_L sequence, we concentrated our work on the

A B C D E F G H I



Fig. 1. Hybridization of B1-8 V_H cDNA to poly(A)⁺ RNA. 5 μ g of poly(A)⁺ RNA from each of the cell lines was analyzed as described in Materials and methods. Lane A, X63.Ag8.653; B, B1-8; C, B1-8.61.V1; D, 251.5; E, 303.20; F, 15.3; G, 260.2; H, 141.3; I, 205.12.

heavy chain variable regions. To determine whether the V_H regions of these proteins are encoded by similar genes we prepared poly(A)⁺ RNA from 10 Ac38-positive hybridoma cell lines and examined them by Northern analysis for hybridization to a cDNA probe from the expressed V_H gene of the B1-8 cell line (see Materials and methods). This 254-bp DNA, encoding amino acids 4–90, is identical in sequence to the germ line gene 186.2 (Bothwell *et al.*, 1981). As shown in Figure 1, poly(A)⁺ RNA from four γ 1-producing cell lines (lanes D, F, G, H), one μ -producing cell line (lane E) and one γ 2b-producing cell line (lane I) hybridizes to the 186.2 gene. Four additional γ 1-producing cell lines (225.4, 214.8, 103.10 and 8.4) also produce mRNA which is homologous to 186.2 (data not shown). X63.Ag8.653, the myeloma partner in the production of these cell lines does not produce homologous RNA (lane A). The B1-8 V_H probe also hybridizes to RNA from the cell line B1-8 (lane B) and from a δ -positive cell line

THIS PAGE BLANK (USPTO)

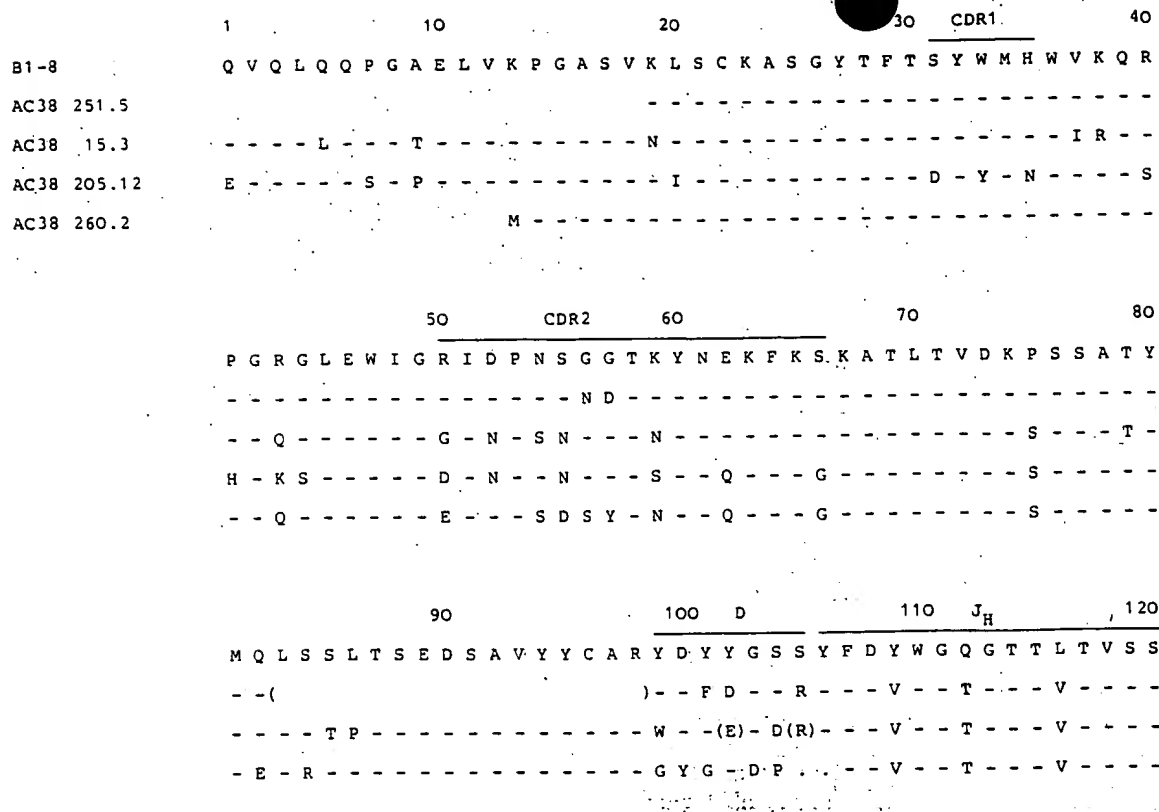


Fig. 2. Protein sequences of the heavy chain variable regions of four Ac38-positive hybridomas. Residues identical to the B1-8 sequence are indicated with dashes.

B1-8.δ1.V1 (lane C), which produces a somatic variant of the B1-8 protein (see Discussion).

We cannot tell by this hybridization analysis how related the different V_H genes expressed in the various hybridomas are to 186.2. Differences in the intensity of hybridization may be partly due to the varying amounts of specific mRNA which these cell lines produce.

Amino acid sequence analysis of V_H regions

To determine how related the different V_H regions of the Ac38-positive proteins are to B1-8 and to each other, we determined the sequences of the V_H regions of four of these proteins: 15.3 ($\gamma 1$), 205.12 ($\gamma 2b$), 260.2 ($\gamma 1$) and 251.5 ($\gamma 1$, NP⁺). The purified antibodies were chemically cleaved with CNBr and the cleavage products size fractionated on Sephadex G-50 as described in Materials and methods. Eluted V_H fragments were either pure enough for direct sequencing or were re-chromatographed after complete reduction and alkylation of disulphide bridges. The protein sequences of the four V_H regions were established by automated Edman degradation and compositional amino acid analysis of two dimensionally separated tryptic and/or thermolytic peptides (Dildrop *et al.*, 1982).

Figure 2 shows the V_H sequences below that of antibody B1-8 (IgM, $\lambda 1$, NP⁺), the protein against which the anti-idiotypic antibody Ac38 had been raised. The homology of the four sequences to the B1-8 sequence and to each other ranges from 72 to 97% in the V_H -gene segments (codon 1–98). The D segments of the antibodies differ from the D segment of B1-8 which belongs to the D_{FL16} group (Kurosawa and Tonegawa, 1982). However, the 15.3 and 251.5 D-segments may also belong to this group, while the 205.12 sequence is an example of a D_{SP2} -like D-segment (Kurosawa

and Tonegawa, 1982). The three J_H segments which have been sequenced here are all J_{H1} whereas B1-8 uses J_{H2} . There is very little variation with respect to the size of the third hypervariable region. The length of two of the D segments is seven amino acids, as in the case of the B1-8 D segment, while the 205.12 D segment contains 6 amino acids. There is also no variation at the D-J border and in the length of the J-segment between the sequence of 15.3, 251.5 and B1-8. However, the 205.12 heavy chain is one amino acid shorter than the others at the D-J junction.

Comparison of V_H regions with related sequences

As mentioned above, the V_H segments of the four Ac38-positive antibodies share striking homology (>70%) with the B1-8 prototype sequence and among themselves. We wished to know whether the differences between these sequences were caused by somatic mutation of the 186.2 (B1-8) gene or whether each of the proteins is encoded by other different, yet related V_H genes. Accordingly, we compared these sequences with previously published V_H gene sequences which also share $\geq 70\%$ homology with the 186.2 gene.

The complete collection of related sequences, presented in Figure 3, contains >40 members. About half of the sequences are germ line sequences from unrearranged V_H genes while the other half represents sequences of expressed rearranged V_H genes. The sequences are ordered by maximizing homology between neighbours and forming a general gradient of decreasing homology from 186.2. The four Ac38-positive antibody sequences (marked by closed arrows) have been placed between their closest relatives. Each of the four sequences fits best to a different sequence on the list: 251.5 contains only two exchanges from 186.2, 15.3 is different from 23 at 11 positions, 260.2 and 124 differ at one position

THIS PAGE BLANK (USPTO)

	1	10	20	30	40	50	60	70	80	90
NP.b 186.2	OVQIQDPGATLVKPGASVKLSCKASGYTFTSYWHIL	LVKQKHQKLEWIGRIID	NSGQTKYHKEFKQKATLTVDKIPSTAYFLLSRLTSDQAVYCAH							
NP.b 81-8										
AC38 25.1.5										
NP.b 145			S							
GAT 117-170.2										
NP.b 106.1										
NP.b 543										
NP.a 3										
NP.a 20										
NP.b 23										
AC38 15.3										
NP.a 33										
NP.b 6										
MPC11 124										
AC38 260.2										
NP.b 102										
NP.b 3										
NP.a 5										
NP.a 31										
MPC11 105										
ARS 36-65										
ARS 10CR.11										
ARS 93G7 CR1+ DNA										
ARS 121F6 CR1+										
ARS 16.7										
ARS 93G7 CR1+ PRO										
ARS 124F1 CR1+										
5C1 1F6 CR1+ARS-										
ARS 10CR.7										
MPC11										
MPC11 111										
MPC11 104										
MPC11 100A										
AC38 205.12										
DEX3 MOPC104E										
DEX3 DEX 12										
DEX3 DEX 10										
DEX3 DEX 9										
DEX3 DEX 8										
MPC11 100B										
UC11										
NP.a 17.2.25										
GAT 65 BB 2.2										
GAT 67 AB 2.9										
GAT 68 PAD 3.9										

Fig. 3. Collection of published V_H gene sequences related to the 186.2 gene: sequences are written down in the one-letter-amino-acid-code (Dayhoff, 1972) numbered according to Kabat *et al.*, 1976. Stop codons are indicated by asterisks, gaps at the a, b, c positions with points, the vertical bar in the gene 6 indicates a missing base in the according codon. The sequences are ordered by maximizing homology between neighbours and forming a general gradient of decreasing homology from 186.2. Ac38-positive hybridoma sequences are indicated with arrows. Available DNA sequences are marked with one asterisk, those of germ line genes with two. Whenever possible specificities are given in front of the name of the sequences. NP.a/NP.b is used for all sequences whose genes have been isolated by using the V_H NP^b DNA probe (S43 derived, Bothwell *et al.*, 1981) as for the 186.2 gene. The only NP-binders among these sequences are B1-8, S43 and 17.2.25. NP.a/NP.b discriminates between BALB/c and C57BL/6 origin. ARS 10CR.11/10CR.7 are the sequences of germ line genes isolated using a 36-65 derived probe (Siekevitz *et al.*, 1983). MPC 11 is used for sequences whose genes were isolated using a MPC 11-derived probe (Givol *et al.*, 1981). Specificities are: Ac38 = monoclonal antibody Ac38, which reacts with many NP^b-positive proteins, was used as immunizing antigen; ARS = anti-*p*-azophenylarsonate, DEX3 = anti- α -1,3 dextran, GAT = anti-poly (Glu 60, Ala 30, Tyr 10), NP = anti-(4'-hydroxy-3-nitro-phenyl)acetyl, 5C1 = rat monoclonal antibody 5C1, which reacts with many CR1⁺ proteins, was used as immunizing antigen. Sequences were taken from following references: NP.b sequences, Bothwell *et al.*, 1981; NP.a sequences, Loh *et al.*, 1983; MPC 11 sequences, Givol *et al.*, 1981; Cohen *et al.*, 1982; Cohen and Givol, 1983. ARS sequences, Estess *et al.*, 1980; Siegelman *et al.*, 1981; Sims *et al.*, 1982; Siekevitz *et al.*, 1983. IF6 CR1⁺, Margolies *et al.*, 1983a; DEX3 sequences, Kehry *et al.*, 1979; Schilling *et al.*, 1980; Tonelle *et al.*, 1981; BCL1, Knapp *et al.*, 1982; GAT sequences, Rocca-Serra *et al.*, 1983a; 1983b. B1-8.V1/V2, Dildrop *et al.*, 1982.

and 205.12 differs from MOPC104E at four positions. The Ac38-positive antibodies therefore probably originate from different V_H genes. Their sequences may derive either directly from the one pointed out here or from other close relatives of these (see Discussion for details).

Discussion

We have previously examined the response of various mouse strains to the monoclonal antibody Ac38 which had been produced against the germ line encoded, NP-binding antibody B1-8 (Reth *et al.*, 1979; Takemori *et al.*, 1982; Nishikawa *et al.*, 1983; Tesch *et al.*, 1983). When antibody Ac38 is used to immunize C57BL/6 mice, from which antibody B1-8 originates, a very large, almost exclusively λ 1-bearing, response is elicited (Tesch *et al.*, 1983). Only 10–20% of these antibodies are NP binding, however. Most of these antibodies also do not bear the other idiotopes associated with B1-8 (Takemori *et al.*, 1982). The vigorous response to antibody Ac38 can be explained by the high frequency (1×10^{-3}) at

which cells expressing antibodies phenotypically similar to those produced in the response are found among LPS-reactive B cells from the spleens of adult and newborn animals and in LPS-reactive B cells generated from bone marrow pre-B cells *in vitro* (Nishikawa *et al.*, 1983). We hypothesized that both the large number of Ac38-bearing B cells and the heterogeneity of the antibody they produce with respect to hapten-binding and the presence of other idiotopes could be due to the existence of many germ-line V_H genes which can encode an Ac38 determinant when in association with a λ 1 light chain. Only a few of these genes would produce antibodies which bind NP or the other B1-8 anti-idiotopes. Additional support for this hypothesis comes from the response to antibody Ac38 of BALB/c and CBA mice, which do not produce idiotope Ac38 in response to NP. Surprisingly, a large λ 1-positive response is elicited in both cases, although none of these antibodies bind NP (Takemori *et al.*, 1982). (A small amount of another B1-8 idiotope, A6-24, is also elicited in CBA mice). In these strains, we would postulate that while they lack the B1-8 V_H gene (186.2), they

THIS PAGE BLANK (USPTO)

contain other members of the large family of related, Ac38-permissive V_H genes. Another result which supports the idea of multiple Ac38-encoding V_H genes is the sequence of a somatic recombinant from B1-8 (B1-8.61.V1) which contains a V_H gene which has replaced most of the 186.2 gene with a portion of a related gene, 102 (Dildrop *et al.*, 1982; Krawinkel *et al.*, 1983). The protein produced by this cell line retains the Ac38 determinant.

An alternative explanation for the large number of Ac38-positive cells could be that the 186.2 gene, which may encode up to 90% of the primary response to NP in C57BL/6, has a higher probability than other V_H genes of recombining to form a complete heavy chain gene. However, only some of the 186.2-D-J combinations may be able to mediate NP binding.

In the current publication, we have chosen to examine the Ac38-encoding V_H gene repertoire directly by analysis of antibodies from λ 1-positive hybridomas produced after immunization with Ac38. Serological analysis of these purified proteins agrees with the earlier results from the analysis of immune sera: only two of the 10 Ac38⁺ proteins examined (251.5 and 303.20) bind to NP. One of these proteins, 251.5, also contains two of the other B1-8 associated idiotopes. Three additional proteins are cross-reactive with one of the anti-B1-8-idiotopes, antibody A6-24, whose target idiotope is also found on NP-non-binding Ac38-positive antibodies elicited in CBA mice (Takemori *et al.*, 1982). The other anti-idiotopes tested do not react with the NP-non-binding hybridoma proteins. In spite of the lack of reactivity observed in the serological analysis, heavy chain mRNAs from the 10 randomly-selected Ac38⁺ hybridomas hybridized to a B1-8 V_H -probe in Northern analysis. The finding that all of the λ 1-bearing, Ac38-positive hybridomas which we have examined are related to 186.2 by hybridization analysis bears out our initial assumption that the λ 1-bearing antibodies would be reactive with Ac38 because they are truly idiotypic (Ab3 β ; Jerne *et al.*, 1982) and not anti-anti-idiotypic (Ab3 α).

Amino acid sequence analysis of four of these heavy chains showed that while they are all related to B1-8 (and to the 186.2 gene) as expected, only one of the proteins, 251.5, appears to be encoded by the 186.2 gene. The sequence of this protein, which is both NP binding and reactive with three anti-idiotope antibodies, differs from B1-8 at only two positions in the V_H region, both in CDR-2. It would appear that the V_H region of this protein is encoded by a somatically mutated 186.2 gene.

The other three heavy chain V_H regions, although in general also similar to 186.2, all possess striking differences from 186.2 and from each other. Do they in fact originate from different V_H genes in the germ line? We believe that they do for two reasons. The first reason involves the pattern of amino acid exchanges which we find: each of the three sequences differs from B1-8 at many positions: 15.3 at 15 out of 98 amino acids, 260.2 at 11 out of 69 amino acids and 205.12 at 20 out of 98 amino acids. However, most of these changes from 186.2 are shared by other related V_H genes which have been sequenced (Figure 3) (15.3 at eight out of the 15 differences, 260.2 at 10 of the 11 differences and 205.12 at all 20 of the differences). In contrast, S43, which is a somatic mutant of 186.2 (Bothwell *et al.*, 1981) differs from 186.2 at seven positions. None of these changes are shared by other members of the gene family. We believe that the 'shared changes' are typical of changes between related germline

genes, while the unique substitutions may be the result of somatic mutations. An additional, perhaps stronger, argument for a distinct V_H gene origin can be made for proteins 205.12 and 260.2. When a search is made for the closest relative of these two proteins, two BALB/c sequences are found. The heavy chain V_H region of 205.12 differs from that of MOPC104E by only four amino acids. The protein sequence of MOPC104E is probably fairly close to its germline-encoded sequence (Schilling *et al.*, 1980). The closest relative of the 260.2 sequence is that of a member of the MPC11 family isolated using an MPC11 V_H gene probe (Cohen and Givol, 1983). This sequence, 124, differs from 260.2 at only one of the 67 positions which have been sequenced. Since it is unlikely that two BALB/c genes could be so nearly reproduced in C57BL/6 by somatic mutation of the 186.2 gene, we feel that it is likely that they represent genes separate from 186.2 and from each other. Although it is difficult to describe alleles in discussing protein sequences from a large family of related genes such as the V_H gene cluster, it appears probable that we have identified in the C57BL/6 strain the closest relatives for two BALB/c V_H genes.

In conclusion, the above results indicate that the V_H regions expressed in antibodies carrying the Ac38 determinant originate from at least four different V_H genes in the germ line. A fifth gene (102; Bothwell *et al.*, 1981) may also be involved in this response since the analysis of an Ac38-positive somatic recombinant of B1-8 contains a V_H region most of which is encoded by that gene (Dildrop *et al.*, 1982; Krawinkel *et al.*, 1983).

It is striking that although these proteins are only 80–85% homologous with B1-8, they still react with the B1-8 anti-idiotope Ac38. Can we use these protein sequences to help define the sequence requirements for the Ac38 idiotope? Previous sequence analysis of NP-binding hybridoma proteins has shown that the Ac38 determinant can be found when the 186.2-encoded V_H is associated with at least three different J_H segments and three different D segments. However, a single amino acid substitution in one of those D segments can lead to the loss of the Ac38 determinant. (For review, see Rajewsky and Takemori, 1983; Zaiss and Beyreuther, unpublished data). This confusing picture is not simplified by considering the present sequences. Although the proteins considered here all utilize J_H1 , previously sequenced Ac38⁺ proteins utilized J_H2 and J_H4 as well. In addition, although one particular mutation in D was found to affect the Ac38 determinant, the sequences presented here contain strikingly different D segments. It is possible that the Ac38 determinant can only be formed when certain combinations of VDJ are made: a particular V may require a particular D or J. Thus, sequence comparisons are not the method of choice for the localization of idiotypic determinants on the V region surface. Similar difficulties have also been found when attempts were made to assign particular idiotopes to hybridoma proteins reactive with dextran (Clevinger, 1981). In that case, although the proteins only differed from each other by the two amino acids in the D regions, it was difficult to predict from those sequences what the idiotypic reactivity of the proteins would be.

Although we cannot define the sequence requirements for the Ac38 idiotope, it is clear that the Ac38 idiotope is a marker for a large family of related V_H genes. These genes can be as little as 72% homologous in their amino acid sequences and can include sequences normally considered part

THIS PAGE BLANK (USPTO)

of the MOPC104E and MPC11 families (Figure 3). In contrast, hybridoma proteins produced after anti-idiotypic immunization in the CRI system of A/J mice were much more restricted (Wysocki and Sato, 1981). Although the antibodies produced could also be either antigen-binding or antigen-non-binding, only the products of a single germ line V_H gene were found (Margolies et al., 1983b; Siekevitz et al., 1983). Results from a variety of other experimental systems also indicate that certain idiotypic determinants can be found on antibodies of different antigen-binding specificities (for review, see Rajewsky and Takemori, 1983). The results presented here show that these idiotypic determinants can be expressed by antibodies whose heavy chain V_H regions are encoded by a family of related V_H genes.

Although we cannot quantitate the number of Ac38-permissible V_H genes in the C57BL/6 genome the number may be quite large. It is probable that the high frequency of Ac38⁺-positive B cells among LPS-reactive cells (1/1000) and the vigorous response of C57BL/6 mice to antibody Ac38 (~500 µg of Ac38+λ1⁺ antibody per ml serum) can be explained by the large number of A38-permissible V_H genes. To a first approximation it therefore appears that the 186.2 gene which dominates the anti-NP response in C57BL/6 mice is not preferentially rearranged in pre-B cells. The cells expressing that gene may be selected at later stages of differentiation.

Materials and methods

Monoclonal antibodies and antigens

The monoclonal antibodies B1-8, Ls136, Ac38, Ac146, As79, A6-24, A25-9, A39-40 and A31-90 have been described previously, including their purification (Reth et al., 1978, 1979; Rajewsky et al., 1981). Keyhole-limpet hemocyanin (KLH) was purchased from Calbiochem, San Diego, CA and bovine serum albumin (BSA) from Behringwerke AG, Marburg/Lahn, FRG. NIP-BSA and Ac38-KLH conjugates were prepared as described previously (Imanishi and Mäkelä, 1973; Takemori et al., 1982).

Immunizations and fusions

3–6-month old C57BL/6 mice, bred in our own colony, were immunized by i.p. injection of 75 µg Ac38-KLH conjugate in complete Freund's adjuvant as described previously (Takemori et al., 1982). Spleen cells from these mice were fused to the myeloma line X63.Ag8.653 (Kearney et al., 1979) either on day 3 (fusion no. 5) or day 6 (fusion no. 7) after priming. In the case of fusion no. 6 the mice were given a secondary injection of antigen in Freund's complete adjuvant 1 month after priming, and the cells were fused on day 3 after the second injection. Fusions and subsequent cloning of hybridoma cells were done as described by Lemke et al. (1978). Hybridoma cells secreting λ1 chain-bearing, Ac38-binding antibodies were detected by a radioimmunoassay (see below). Cloned hybridoma lines were propagated in the peritoneum of pristane-pre-treated (BALB/cx C57BL/6)F₁ mice bred in our colony. Ascites fluid produced by these mice was analyzed for the presence of monoclonal antibody by microzone electrophoresis and serologically. Hybridoma cells were analyzed for the production of intracellular Ig by fluorescence microscopy (Kearney et al., 1979), using fluorescein-coupled anti-idiotypic or anti-Ig antibodies of appropriate specificity (gift of Dr. A. Radbruch).

Serology

A two-stage radioimmunoassay was used for the determination of λ1 chain-bearing antibodies with specificity for the various anti-idiotypes and/or for the NIP hapten (Reth et al., 1979). Briefly, wells of 96-well plastic plates (Dynatech Deutschland GmbH, Plochingen, FRG) were coated with one of the anti-idiotypic antibodies or with NIP₄-BSA at a concentration of 10–20 µg/ml. Serial dilutions of culture supernatant or ascites fluid containing the antibody to be determined were added and, after overnight incubation, bound λ1 chain-bearing antibodies were detected by radiolabelled antibody Ls136.

Isolation of poly(A)⁺ RNA

Isolation of total RNA and subsequent enrichment of the poly(A)⁺ fractions were performed as described in Siekevitz et al. (1983).

Preparation of nucleic acid probe

pABμ-11, a cDNA clone containing the expressed V_H gene from the cell line B1-8, has been described previously (Bothwell et al., 1981). After *Pst*I and

*Hinf*I restriction endonuclease digestion, a 254-bp fragment was purified from this clone according to the methods described by Maniatis et al. (1982).

Transfer of poly(A)⁺ RNA to solid support

Poly(A)⁺ RNA was glyoxylated and run on 1.2% agarose gels as described by McMaster and Carmichael (1977). RNA was transferred to nitrocellulose paper as described by Thomas (1980).

Hybridization of radiolabelled nucleic acids

Double-stranded DNA was radiolabelled by nick translation with DNA polymerase I according to the method of Rigby et al. (1977). Hybridization of radiolabelled DNA to RNA bound to nitrocellulose was as described by Thomas (1980). Autoradiography was performed at –70°C using Kodak XAR-5 film and a DuPont Kronex Lightning-plus intensifying screen.

Sequence analysis

Antibodies were purified from ascites fluid by (NH₄)₂SO₄ precipitation followed by ion-exchange chromatography on DEAE-cellulose (Reth et al., 1979) or in the case of NP-binding antibody 251.5 by absorption to (NIP-cap)-Sephacrose and subsequent elution from the washed sorbent with 10^{–3} M NIP-cap [(5-iodo-3-nitro-phenyl)acetyl-caproic acid] in phosphate-buffered saline (Brüggemann et al., 1982).

The isolation of heavy chain variable region CNBr fragments and their subsequent sequence analysis was performed as described previously (Dildrop et al., 1982).

Acknowledgements

We are grateful to Ms. G. von Hesberg, Ms. G. Zimmer and Mr. K. Neifer for expert technical help. This work was supported by the Deutsche Forschungsgemeinschaft through SFB 74, the Minister for Science and Research of NRW and the Fazit Foundation. M.S. is a fellow of The Jane Coffin Childs Memorial Fund for Medical Research.

References

- Bothwell, A.L.M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewsky, K. and Baltimore, D. (1981) *Cell*, **24**, 625–637.
- Bothwell, A.L.M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewsky, K. and Baltimore, D. (1982) *Nature*, **298**, 380–382.
- Brüggemann, M., Radbruch, A. and Rajewsky, K. (1982) *EMBO J.*, **1**, 629–634.
- Clevinger, B. (1981) in Westen-Schnorr, I. (ed.), *Idiotypes: Antigens on the Inside*, Proceedings of the Workshop at the Basel Institute for Immunology, Nov. 1981, Editions Roche, Basel, pp. 77–83.
- Cohen, J.B., Effron, K., Rechari, G., Ben-Neriah, Y., Zakut, R. and Givol, D. (1982) *Nucleic Acids Res.*, **10**, 3353–3370.
- Cohen, J.B. and Givol, D. (1983) *EMBO J.*, **2**, 2013–2018.
- Dayhoff, M. (1972) *Atlas of Protein Sequence and Structure*, Vol. 5, D2, published by National Biomedical Research Foundation, Washington.
- Dildrop, R., Brüggemann, M., Radbruch, A., Rajewsky, K. and Beyreuther, K. (1982) *EMBO J.*, **1**, 635–640.
- Estess, P., McCumber, L., Siegelman, M., Waldschmidt, T., Pacifico, A. and Capra, J.D. (1980) *Progr. Immunol.*, **4**, 3–16.
- Givol, D., Zakut, R., Effron, K., Rechari, G., Ram, D. and Cohen, J.B. (1981) *Nature*, **292**, 426–430.
- Imanishi, T. and Mäkelä, O. (1973) *Eur. J. Immunol.*, **3**, 323–330.
- Jerne, N.K., Roland, J. and Cazenave, P.-A. (1982) *EMBO J.*, **1**, 243–247.
- Kabat, E.A., Wu, T.T. and Bilofsky, H., eds. (1976) *Variable Regions of Immunoglobulin Chains*, Medical Computer Systems, Bolt Beranek and Newman Inc., Cambridge, MA.
- Kearney, J.F., Radbruch, A., Liesegang, B. and Rajewsky, K. (1979) *J. Immunol.*, **123**, 1548–1550.
- Kehry, M., Sibley, C., Fuhrman, J., Schilling, J. and Hood, L.E. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 2932–2936.
- Knapp, M.R., Liu, C.P., Newell, N., Ward, R.B., Tucker, P.W., Strober, S. and Blattner, F. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2996–3000.
- Krawinkel, U., Zobelein, G., Brüggemann, M., Radbruch, A. and Rajewsky, K. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4997–5001.
- Kurosawa, Y. and Tonegawa, S. (1982) *J. Exp. Med.*, **155**, 201–218.
- Lemke, H., Hammerling, G.J., Höhmman, C. and Rajewsky, K. (1978) *Nature*, **271**, 249–251.
- Loh, D.Y., Bothwell, A.L.M., White-Scharf, M.E., Imanishi-Kari, T. and Baltimore, D. (1983) *Cell*, **33**, 85–93.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning*, published by Cold Spring Harbor Laboratory Press, NY.
- Margolies, M.N., Juszczak, E.C., Near, R., Marshak-Rothstein, A., Rothstein, T.L., Sato, V.L., Siekevitz, M., Smith, J.A., Wysocki, L.J. and Gelfand, M.L. (1983a) *Ann. N.Y. Acad. Sci.*, in press.

THIS PAGE BLANK (USPTO)

- Margolies, M.N., Wysocki, L.J. and Sato, V.L. (1981) *J. Immunol.*, **130**, 515-518.
- McMaster, G.K. and Carmichael, G.G. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 4835-4838.
- Nishikawa, S., Takemori, T. and Rajewsky, K. (1983) *Eur. J. Immunol.*, **13**, 318-325.
- Rajewsky, K., Takemori, T. and Reth, M. (1981) in Hämmerling, G.J., Hämmerling, U. and Kearney, J.F. (eds.), *Monoclonal Antibody and T Cell Hybridoma: Perspective and Technical Advances*, Elsevier/North-Holland Biomedical Press, Amsterdam, pp. 399-409.
- Rajewsky, K. and Takemori, T. (1983) *Annu. Rev. Immunol.*, **1**, 569-607.
- Reth, M., Hämmerling, G.J. and Rajewsky, K. (1978) *Eur. J. Immunol.*, **8**, 393-400.
- Reth, M., Imanishi-Kari, T. and Rajewsky, K. (1979) *Eur. J. Immunol.*, **9**, 1004-1013.
- Rigby, W.J.B., Dieckermann, M., Rhodes, C. and Berg, P. (1977) *J. Mol. Biol.*, **113**, 237-251.
- Rocca-Serra, J., Matthes, H.W., Kaartinen, M., Milstein, C., Theze, J. and Fougereau, M. (1983a) *EMBO J.*, **2**, 867-872.
- Rocca-Serra, J., Tonelle, C. and Fougereau, M. (1983b) *Nature*, **304**, 353-355.
- Schilling, J., Clevinger, B., Davie, J.M. and Hood, L. (1980) *Nature*, **283**, 35-40.
- Siegelman, M., Slaughter, C., McCumber, L., Estess, P. and Capra, J.D. (1981) in Janeway, C., Sercarz, E.E. and Wigzell, H. (eds.), *Immunoglobulin Idiotypes (ICN-UCLA Symposia on Molecular and Cellular Biology)* Vol. 20, Academic Press, Inc., NY, pp. 135-158.
- Siekevitz, M., Huang, S.Y. and Geffer, M.L. (1983) *Eur. J. Immunol.*, **2**, 123-132.
- Sims, J., Rabbitts, T.H., Estess, P., Slaughter, C., Tucker, P.W. and Capra, J.D. (1982) *Science (Wash.)*, **216**, 309-311.
- Takemori, T. and Rajewsky, K. (1981) *Eur. J. Immunol.*, **11**, 618-625.
- Takemori, T., Tesch, H., Reth, M. and Rajewsky, K. (1982) *Eur. J. Immunol.*, **12**, 1040-1046.
- Tesch, H., Takemori, T. and Rajewsky, K. (1983) *Eur. J. Immunol.*, in press.
- Thomas, P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201-5205.
- Tonelle, C., Pierres, M., Ju, S.T., Moinier, D. and Fougereau, M. (1981) *Mol. Immunol.*, **18**, 979-984.
- Weigert, M. and Riblet, R. (1976) *Cold Spring Harbor Symp. Quant. Biol.*, **41**, 837-846.
- Wysocki, L.J. and Sato, V.L. (1981) *Eur. J. Immunol.*, **11**, 832-839.

Received on 21 October 1983; revised on 12 December 1983

THIS PAGE BLANK (USPTO)

Volume 2 Number 3 March 1984

STANFORD UNIVERSITY
LIBRARIES

THE EMBO JOURNAL

Published monthly for
the European Molecular Biology Organization
by IRL Press

ISSN 0261 4189 Coden EMJODG

THIS PAGE BLANK (USPTO)



XP 000069445

C12N15/82

mechanism of NA release from rat cerebral-cortex synaptosomes. Because immunocytochemical studies have shown that B-50 is found in synapses throughout the brain^{21,22}, B-50 could be more generally involved in transmitter release. Three lines of evidence suggest that the phosphorylation of B-50 by PKC is essential for stimulus-secretion coupling during transmitter release: (1) phorbol esters that directly activate PKC enhance the release of a variety of neurotransmitters¹⁻⁴; (2) by using an antibody-independent approach, we have previously shown that depolarization-induced neurotransmitter release from non-permeabilized synaptosomes and hippocampal slices is closely correlated with a PKC-mediated increase in B-50 phosphorylation^{18,23}; and (3) here we have shown that anti-B-50 IgG inhibits B-50 phosphorylation as well as Ca²⁺-dependent transmitter release. If B-50 phosphorylation by PKC is indeed involved in the mechanism of transmitter release, then a long-term increase in PKC-mediated B-50 phosphorylation^{7,17} could be one of the mechanisms underlying the increase in the release of glutamate that occurs during long-term potentiation^{5-7,24,25}.

In view of the localization of B-50 at the inner leaflet of the plasma membrane^{21,26}, we suggest that B-50 is involved in the regulation of vesicle fusion with the plasma membrane, a process in which the vesicle-associated protein synapsin I (a substrate of calmodulin-dependent kinases) has also been implicated^{27,28}. But the difference in the localization of phosphorylating enzymes of these two proteins indicates that they have distinct roles in the transmitter release process. It may be that the regulatory role of B-50 in vesicle fusion is not limited to transmitter release, but extends to membrane-fusion processes during neurite outgrowth^{29,30}. It remains to be investigated to what extent calmodulin binding¹⁴ and modulation of phosphatidylinositol 4-phosphate kinase activity^{12,13}—putative properties of B-50—are also involved in controlling neurotransmitter release. □

Received 17 April; accepted 11 September 1989.

1. Allgaier, C., Von Kügelgen, C. & Rüttung, G. *Eur. J. Pharmacol.* **129**, 389-392 (1986).
2. Malenka, R. C., Aoyou, G. S. & Nicoll, R. A. *Brain Res.* **403**, 198-203 (1987).
3. Nichols, R. A., Haycock, J. W., Wang, J. K. T. & Greengard, P. *J. Neurochem.* **48**, 615-621 (1987).
4. Kaczmarek, L. K. *Trends Neurosci.* **10**, 30-34 (1987).
5. Malenka, R. C., Madison, D. V. & Nicoll, R. A. *Nature*, **321**, 175-177 (1986).
6. Malinow, R., Madison, D. V. & Tsien, R. W. *Nature* **335**, 820-824 (1988).
7. Akeris, R. F., Lovinger, D. M., Colley, P. A., Linden, D. J. & Routenberg, A. *Science* **231**, 587-589 (1986).
8. Zwiers, H., Schotman, P. & Gispen, W. H. *J. Neurochem.* **34**, 1689-1699 (1980).
9. Aoyou, V. J., Zwiers, H. & Gispen, W. H. *J. Neurochem.* **41**, 649-653 (1983).
10. De Graan, P. N. E., Dekker, L. V., Oestreicher, A. B., Van der Voorn, L. & Gispen, W. H. *J. Neurochem.* **52**, 17-23 (1989).
11. Skene, J. H. P. *Rev. Neurosci.* **12**, 127-156 (1989).
12. Jüttles, J. et al. *Nature* **296**, 623-625 (1980).
13. Gispen, W. H., Van Ongen, C. J., De Graan, P. N. E., Oestreicher, A. B. & Zwiers, H. in *Inositol and Phosphoinositides* (eds Gledhill, J. E., Hauser, G. & Eichberg, J.) 399-413 (Humana, Clifton, New Jersey 1985).
14. Andreassen, T. J., Luetje, C. W., Heideman, W. & Storm, D. R. *Biochemistry* **22**, 4615-4618 (1983).
15. Skene, J. H. P. & Willard, M. *J. Cell Biol.* **89**, 86-95 (1981).
16. Goslin, K., Schreyer, D. J., Skene, J. H. P. & Banker, G. *Nature* **336**, 672-674 (1988).
17. Lovinger, D. M. et al. *Brain Res.* **343**, 137-143 (1985).
18. Dekker, L. V., De Graan, P. N. E., Versteeg, D. H. G., Oestreicher, A. B. & Gispen, W. H. *J. Neurochem.* **52**, 24-30 (1989).
19. Howell, T. W. & Gomperts, B. D. *Biochim. biophys. Acta* **927**, 177-183 (1987).
20. Vernage, M., Besselsen, E., Lopes, F. H. & Ghijzen, W. E. J. M. *J. Neurochem.* **51**, 1667-1674 (1988).
21. Gispen, W. H., Leunissen, J. L. M., Oestreicher, A. B., Verkleij, A. J. & Zwiers, H. *Brain Res.* **328**, 381-385 (1985).
22. Benowitz, L. L., Apostolides, P. J., Perrone-Bizzozzi, N., Finkbeiner, S. P. & Zwiers, H. *J. Neurosci.* **8**, 339-352 (1988).
23. Schram, L. H. et al. *Soc. Neurosci. Abstr.* **14**, 197.15 (1988).
24. Oalpin, A. C., Errington, M. L. & Bliss, T. V. P. *Nature* **297**, 496-498 (1989).
25. Anikstein, L., Rosen, M. P., Amzel, R. & Ben-Ari, Y. *Neuroscience* **28**, 387-392 (1989).
26. Skene, J. H. P. & Verag, L. *J. Cell Biol.* **108**, 613-624 (1989).
27. Böhler, M. & Greengard, P. *Nature* **328**, 704-707 (1987).
28. Hirokawa, N., Sobue, K., Kanda, K., Harada, A. & Yanfu, H. *J. Cell Biol.* **108**, 111-126 (1989).
29. van Hoof, C. Q. M. et al. *J. Cell Biol.* **108**, 1115-1125 (1989).
30. Zuber, M. X., Goodman, O. W., Karns, L. R. & Fishman, M. C. *Science* **244**, 1193-1195 (1989).
31. De Langen, C. D. J., Hogenboom, F. & Mulder, A. H. *Eur. J. Pharmacol.* **60**, 79-89 (1979).
32. Bergmeyer, H. U. & Bernt, E. in *Methoden der Enzymatischen Analyse* (ed. Bergmeyer, H. U.) 533-538 (Verlag Chemie, Weinheim, 1970).
33. Oestreicher, A. B., van Ongen, C. J., Zwiers, H. & Gispen, W. H. *J. Neurochem.* **41**, 331-340 (1983).
34. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. Biol. Chem.* **193**, 265-275 (1951).

ACKNOWLEDGMENTS. We thank W. M. Pijnappel for experimental assistance, Dr W. J. Vaartjes for advice on the lactate dehydrogenase assay and E. Kluis and L. Claessens for preparing the manuscript.

Production of antibodies in transgenic plants *p. 76-78*

Andrew Hiatt, Robert Cafferkey & Katherine Bowdish

Department of Molecular Biology, The Research Institute of Scripps Clinic, 10566 North Torrey Pines Road, La Jolla, California 92037, USA

COMPLEMENTARY DNAs derived from a mouse hybridoma messenger RNA were used to transform tobacco leaf segments followed by regeneration of mature plants. Plants expressing single gamma or kappa immunoglobulin chains were crossed to yield progeny in which both chains were expressed simultaneously. A functional antibody accumulated to 1.3% of total leaf protein in plants expressing full-length cDNAs containing leader sequences. Specific binding of the antigen recognized by these antibodies was similar to the hybridoma-derived antibody. Transformants having γ - or κ -chain cDNAs without leader sequences gave poor expression of the proteins. The increased abundance of both γ - and κ -chains in transformants expressing assembled gamma-kappa complexes was not reflected in increased mRNA levels. The results demonstrate that production of immunoglobulins and assembly of functional antibodies occurs very efficiently in tobacco. Assembly of subunits by sexual cross might be a generally applicable method for expression of heterologous multimers in plants.

The source of immunoglobulin mRNAs was a hybridoma cell line expressing a catalytic IgG₁ antibody (6D4) which binds a low molecular weight phosphonate ester (P3) and catalyses the hydrolysis of certain carboxylic esters. Constructs used for immunoglobulin expression in plants consisted of coding-length cDNAs of the 6D4 γ - or κ -chain with or without their leader sequences. These cDNAs were modified to contain terminal EcoRI restriction enzyme digestion sites and were ligated into the constitutive plant expression vector pMON530 (ref. 2) to form pH101 (kappa, no leader), pH102 (kappa, leader), pH201 (gamma, no leader) and pH202 (gamma, leader). We transformed tobacco plants using *Agrobacterium* containing each of these four plasmids³ and screened leaf extracts from regenerated transformants for the presence of immunoglobulin heavy or light chains by enzyme-linked immunosorbent assay (ELISA)⁴. Transformants expressing individual immunoglobulin chains were then sexually crossed to produce progeny expressing both chains. The results of the ELISA revealed high levels of kappa and gamma chains accumulating in individual plants containing DNA from both pH102 and pH202 (Table 1; Fig. 2a). We verified the expression of both heavy and light chains by western blotting (Fig. 1). From the ELISAs, we judged that virtually all the γ - and κ -chains in these plants were assembled into gamma-kappa complexes (Table 1). Western blots provided additional evidence for assembled antibodies in that, under non-reducing conditions, most of the immunoreactive γ - and κ -chains aggregated at a high molecular weight (Fig. 1).

The binding specificity of the assembled gamma-kappa complexes was studied in ELISAs in which a P3-bovine serum albumin conjugate was used as antigen. The antigen binding by antibody derived from plants was equivalent to antigen binding by the 6D4 hybridoma antibody. Incubation of plant extracts or the purified 6D4 antibody with 50 $\mu\text{mol l}^{-1}$ P3 for 3 h at 25 °C before addition to the ELISA eliminated antibody binding to the P3-BSA conjugate, demonstrating that binding was specific for the P3 hapten. Half-maximal inhibition occurred with 10 $\mu\text{mol l}^{-1}$ free P3 for both hybridoma and plant-derived antibodies.

Transformants derived from the leaderless constructs pH101 and pH201 contained very low levels of κ - and γ -chains respectively, but Southern and northern blots (Fig. 2) demonstrated the presence of transforming DNA and immuno-

immunoglobulin chains contained assembled gamma-kappa complexes (Table 1).

The increased recovery of immunoglobulin epitopes from transformants expressing full-length cDNAs was not reflected in increased mRNA transcript levels. Northern blots (Fig. 2b) comparing pHi201 and pHi202 transformants, for example, revealed nearly equivalent levels of heavy-chain transcripts, although ELISAs indicated a 40-fold increase in accumulation of heavy-chain protein in the pHi202 transformant. Likewise, immunoglobulin mRNA levels in a plant producing large amounts of assembled antibodies were not significantly different from the parental plants that accumulated low levels of immunoglobulin chains (Fig. 2b).

Our results show that individual cDNAs for immunoglobulin κ - and γ -chains can be efficiently expressed in tobacco to form functional antibodies. Assembly of immunoglobulin chains by sexual cross in plants represents a useful alternative to the expression by a single vector of both gamma and kappa cDNAs as in yeast or bacteria^{5,7}, or double transformation with vectors containing individual cDNAs^{8,9}. Potentially, this method is applicable to the assembly of oligomers other than antibodies. The characterization of antibodies produced in plants (glycosy-

lation, processing of leader sequences, cytolocalization and turnover) will be described in a later paper.

In B lymphocytes, immunoglobulin processing and assembly occurs in the endoplasmic reticulum/Golgi in a process that may be promoted by heavy-chain binding proteins present in the endoplasmic reticulum^{10,11}. Plant cells may also have a system for multimer assembly in their endoplasmic reticulum/Golgi that can recognize immunoglobulin chains. Alternatively, assembly may occur spontaneously, given sufficient levels of each chain in the appropriate cellular compartment. Our results demonstrate that plants require a signal sequence for efficient assembly of γ - and κ -subunits. The presence of the mouse leader sequence clearly augments the accumulation of individual chains. This might be the result of an enhanced translation of the immunoglobulin messengers or an increased stability of each protein as a result of subcellular sequestering or secretion. The yield of each chain is increased in plants expressing both gamma and kappa, indicating that assembly of the gamma-kappa complex might enhance stability.

TABLE 1 Expression and assembly of immunoglobulin gamma and kappa chains in tobacco

Accumulation of γ - or κ -chains in transformed plants*			
γ NL 30 \pm 16 (60)	γ L 1,412 \pm 270 (2,400)	γ L(κ L) 3,330 \pm 2,000 (12,800)	γ NL(κ NL) 32 \pm 26 (60)
κ NL 1.4 \pm 1.2 (3.5)	κ L 56 \pm 5 (80)	κ L(γ L) 3,700 \pm 2,300 (12,800)	κ NL(γ NL) 6.5 \pm 5 (20)
Distribution and assembly in crosses†			
	γ only	κ only	$\gamma\kappa$ Null
κ NL \times γ NL	4	6	3 (0% assembly) 5
κ L \times γ L	3	10	11 (95 \pm 16% assembly) 4

* Accumulation of individual gamma and kappa chains (in ng per mg total protein) was estimated by ELISA*. Microtitre wells were coated with a goat anti-mouse heavy or light chain-specific IgG (Fisher) in 150 mM NaCl, 20 mM Tris-HCl, pH 8.0 (TBS), followed by blocking with 5% non-fat dry milk in TBS. Plant leaves were homogenized in a mortar and pestle at 4 °C after removal of the midvein. To the supernatant a quarter volume of 5 \times TBS was added, and 50 μ l of 1 in 2 serial dilutions were added to each microtitre well. After 18 h at 4 °C, microtitre wells were washed with distilled water at room temperature. Bound γ - or κ -chains were reacted with goat anti-mouse heavy or light chain-specific antibodies conjugated to horseradish peroxidase for 2 h at 37 °C in TBS, and detected according to the manufacturer's instructions. Control microtitre wells contained extracts from plants transformed with pMON530 vector. Values given as mean (\pm s.d.) are derived from at least two determinations per plant and do not include transformants producing no detectable γ - or κ -chain. At least nine plants were assayed in each category. All values are given as ng per mg of total protein in the extract and are derived from the quantity of purified 6D4 antibody required to give an equivalent colour development in ELISA. Total protein in the extract was determined by the Bio-Rad Coomassie assay. Complementary DNAs containing no leader sequences are referred to as γ NL and κ NL; γ L and κ L refer to cDNAs with leader sequences; $\gamma(\kappa)$ refers to gamma chains in a plant that also expresses κ -chains, and vice versa. Numbers in parentheses are values for plants with the highest levels of accumulation.

† The number of plants expressing γ - or κ -chains among the progeny of a sexual cross. The ELISA for assembly used horseradish peroxidase-conjugated anti- κ -chain-specific antibodies to detect antigen bound to microtitre wells coated with unlabelled anti- γ -chain-specific antibodies, and vice versa. Values derived from these assays were used to calculate the per cent of assembly by comparison with the purified 6D4 antibody. This was determined at least three times for each $\gamma\kappa$ plant. The per cent assembly is expressed in parentheses as the mean \pm s.d.

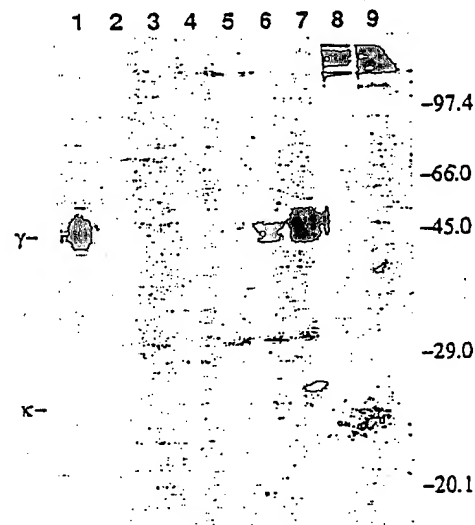


FIG. 1 Western blot of leaf proteins from transgenic tobacco plants expressing immunoglobulin chains. Leaf segments (1 g) from mature plants were homogenized in a mortar and pestle with 1 ml 0.05 M Tris-HCl, pH 7.5, 1 mM phenylmethanesulphonyl fluoride. Extracts were boiled in 4 M urea, 1% SDS, with or without 2 mM dithiothreitol (DTT) as indicated, for 3 min. SDS-PAGE in 10% acrylamide¹⁶ and blotting of the proteins to nitrocellulose¹⁷ were performed as described. Blots were preincubated for 6 h at 4 °C in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.01% Tween 20 (TBST) containing 5% BSA, and 0.5% non-fat dried milk before the addition of antibodies. A biotinylated goat anti-mouse whole IgG antibody (Cappel), diluted 1:500 in TBST was used to probe the blots at 4 °C for 24 h. A variety of commercially available antibodies (anti-mouse IgGs) were used in other experiments with similar results. Antibody binding was visualized after binding of streptavidin-conjugated alkaline phosphatase (25 °C, for 2 h) by incubation in 300 μ g ml⁻¹ nitroblue tetrazolium and 150 μ g ml⁻¹ 5-bromo-4-chloro-3-indolyl phosphate. In lanes 1–7, 40 μ l of each extract containing DTT; lanes 8 and 9, 40 μ l extract without DTT. Lane 1, 100 ng purified antibody from the 6D4 hybridoma; lane 2, 15 μ g wild-type plant-extract protein; lane 3, 15 μ g protein from a plant transformed with truncated κ -chain cDNA (pHi101) containing no leader sequence; lane 4, 15 μ g from plant transformed with truncated γ -chain cDNA (pHi201); lane 5, 15 μ g from a full-length kappa cDNA transformant (pHi102); lane 6, 15 μ g from a full-length γ -chain cDNA transformant (pHi202); lane 7, 15 μ g from an F1 plant derived from the cross between a kappa and a gamma producer; lane 8, 100 ng 6D4 antibody (no DTT); lane 9, same as lane 7, except no DTT. Gamma and kappa on the left refer to the positions of the 6D4 heavy and light chains; positions of molecular weight (given in thousands) markers are shown on the right. By ELISA, extracts in lanes 3–5 contained very low levels of κ - or γ -chains (<0.008% of total protein, Table 1), whereas extracts in lanes 6, 7 and 9 contained 0.24, 1.3 and 1.3% immunoglobulin respectively.

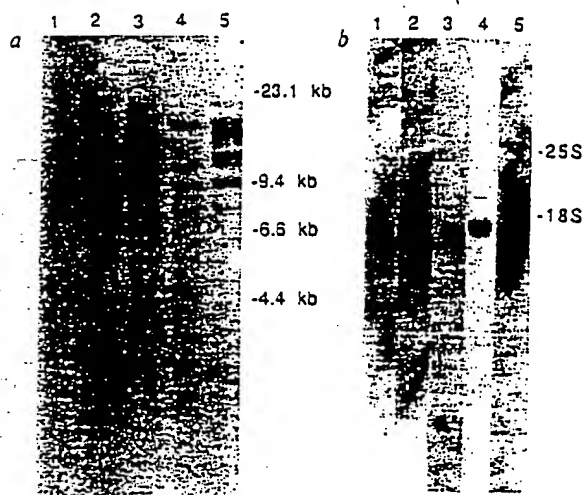


FIG. 2 Southern and northern blots of leaf DNA and RNA from transgenic plants expressing immunoglobulin cDNAs. **a**, DNA was extracted from 1 g mature leaf tissue after freezing the fresh segments in liquid N_2 . Homogenization into urea mix¹⁸ was in a mortar and pestle. The homogenate was extracted with phenol:CHCl₃ (1:1) and the nucleic acids precipitated by addition of one volume of isopropyl alcohol. The resuspended DNA (20 µg each) was cut with *Hind*III and Southern-blotted as described¹⁹. The probe used for both Southern and northern blots was ³²P-labelled pMON530 plasmid² containing either a kappa cDNA or a gamma cDNA. Both cDNAs were used in the hybridization shown. Lane 1, DNA from a transformant expressing a light-chain cDNA without a leader sequence (pHi101); lane 2, DNA from a heavy-chain cDNA transformant, no leader (pHi201); lane 3, DNA from a transformant expressing full-length light chain with leader (pHi102); lane 4, DNA from a transformant expressing heavy chain with leader (pHi202); lane 5, DNA from an F₂ plant derived from a cross between plants expressing full-length gamma or kappa cDNAs (pHi102 × pHi202). **b**, Extraction of RNA from 1 g fresh leaf tissue was by homogenization in 10 ml 0.1 M Tris-HCl, pH 9.0, and 10 ml phenol saturated with this buffer, using a Polytron at high speed. Nucleic acids were precipitated by addition of one tenth volume 3.0 M sodium acetate, pH 5.0, and 1.5 volumes isopropyl alcohol. Resuspended RNA was electrophoresed in gels containing formaldehyde and northern blotted onto nylon membranes (Amersham) as described¹⁹. Lane 1, RNA from a transformant expressing a light-chain cDNA without a leader sequence (pHi101); lane 2, RNA from a heavy-chain cDNA transformant, no leader (pHi201); lane 3, RNA from a transformant expressing full-length light chain with leader (pHi102); lane 4, RNA from a transformant expressing heavy chain with leader (pHi202); lane 5, RNA from an F₂ plant derived from a cross between plant expressing full-length gamma or kappa cDNAs (pHi102 × pHi202). Total RNA (20 µg) was loaded in each lane. Lanes from separate hybridizations were aligned with respect to the 18S (1,900 base pairs) and 25S (3,700 base pairs) ribosomal RNA bands on the blots, detected by methylene-blue staining.

Expression of functional antibodies from transcripts that do not contain signal sequences may require modifications to yield alternative antigen-binding structures (such as single-chain antigen-binding proteins^{6,7}) that do not need to be assembled. Thus binding of constituents of metabolic pathways involved in morphogenesis, stress responses or plant-pathogen interactions could be used to further our understanding of these processes in a way analogous to the blocking by microinjected antibodies of specific protein functions in mammalian cells^{12,13}.

As large macromolecules such as protein multimers do not pass through plant cell walls^{14,15}, the binding by antibodies of small organic molecules (toxins, herbicides, plant hormones, organic chelates, for example) that are permeable to the cell wall might result in the net uptake and retention of these molecules in the plant. Accumulation of functioning antibodies may provide new options for the recovery of an array of environmental contaminants, as well as other biologically significant organics. Catalytic processing of small molecules within cell

wall boundaries by antibodies may also become a generalizable strategy for the elimination or modification of permeable organic compounds and could introduce new catalytic properties in existing metabolic pathways.

Received 21 July; accepted 22 September 1989.

1. Tramontano, A., Janda, K. D. & Lerner, R. A. *Science* **234**, 1565-1570 (1987).
2. Rogers, S. G., Klee, H. J., Horsch, R. B. & Fraley, R. T. *Meth. Enzym.* **153**, 253-277 (1987).
3. Horsch, R. B. et al. *Science* **227**, 1229-1231 (1985).
4. Engvall, E. & Perlmann, P. *J. Immun.* **109**, 129-135 (1972).
5. Carlson, J. R. *Mol. cell. Biol.* **8**, 2638-2646 (1988).
6. Huston, J. S. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5879-5883 (1988).
7. Bird, R. E. et al. *Science* **242**, 423-426 (1988).
8. Wood, C. R. et al. *Nature* **314**, 446-449 (1985).
9. Horwitz, A. H., Chang, C. P., Better, M., Hellstrom, K. E. & Robinson, R. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8678-8682 (1988).
10. Tartakoff, A. & Vassalli, P. *J. Cell Biol.* **83**, 284-299 (1979).
11. Munro, S. & Pelham, H. R. B. *Cell* **46**, 291-300 (1986).
12. Lamo, N. J. C. et al. *J. Cell Biol.* **106**, 1955-1971 (1988).
13. Riazowol, K. T., Vosatka, R. L., Zil, E. B., Lamo, N. J. & Feramisco, J. R. *Mol. cell. Biol.* **8**, 1670-1676 (1988).
14. Milburn, J. A. *Water Flow in Plants* Longman (London, 1979).
15. Carotta, N., Sabatini, O., Montezinos, O. & Omer, D. P. *Science* **205**, 1144-1147 (1979).
16. Chua, N.-H. *Meth. Enzym.* **69**, 434-446 (1980).
17. Harlow, E. & Lane, D. *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1988).
18. Shure, M., Wessler, S. & Fedoroff, N. *Cell* **35**, 225-233 (1983).
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).

ACKNOWLEDGEMENTS. We thank Andrew Gelman for technical assistance, and Drs Mich Hen, Ric Lerner, Richard Farr and Stephen Mayfield for discussion during the preparation of the manuscript.

Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor

David Camerini, Stephen P. James*, Ivan Stamenkovic† & Brian Seed

Departments of Molecular Biology, Massachusetts General Hospital, Genetics and † Pathology, Harvard Medical School, Boston, Massachusetts 02114, USA

* Section on Mucosal Immunity, Laboratory of Clinical Investigation, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892, USA

THE human pan-leukocyte antigen Leu-8 has attracted wide interest because its presence or absence identifies suppressor-inducer and helper-inducer CD4⁺ T-lymphocyte subsets respectively. We report here that Leu-8 is the human homologue of the mouse Mel-14 homing receptor, a molecule that promotes the initial adhesion of blood-borne lymphocytes to the specialized post-capillary endothelium of peripheral lymph nodes. We also show that Leu-8 can adopt both conventional and phospholipid anchored forms, a finding that may have relevance in the context of antigen shedding following activation or homing. The assignment of lymphocytes to different functional classes based on lymph node homing potential may represent a more general association between lymphocyte function and tissue distribution.

Two complementary DNA clones encoding Leu-8 determinants were isolated from a human T-cell library by method previously described^{1,2}. DNA sequence analysis showed that the longer insert of the two contains 2,350 residues, whereas the shorter lacks 436 internal residues but is otherwise identical (Fig. 1). The predicted polypeptide sequences were found to diverge at their C-termini.

The protein encoded by the larger insert bears a strongly hydrophobic putative membrane spanning domain near its C terminus, followed by several positively charged residues resembling a cytoplasmic anchor sequence. The protein is closely related to the recently described murine Mel-14 homing receptor^{3,4} (Fig. 1) and the corresponding cDNA sequence shares

THIS PAGE BLANK (USPTO)

Synthesis and self-assembly of a functional monoclonal antibody in transgenic *Nicotiana tabacum*

Klaus Düring,^{1,*} Sigrun Hippe,² Fritz Kreuzaler^{1,3} and Jeff Schell¹

¹Max Planck Institute for Plant Breeding, D-5000 Köln, FRG; ²Institute for Biology III, RWTH Aachen, D-5100 Aachen, FRG; ³Institute for Biology I, RWTH Aachen, D-5100 Aachen, FRG; ⁴present address: University of Hamburg, Center for Applied Plant Molecular Biology, Ohnhorststr. 18, D-2000 Hamburg 52, FRG (*author for correspondence)

Received 14 November 1989; accepted in revised form 14 May 1990

Key words: monoclonal antibody, multimeric foreign protein, self-assembly, signal peptide, transgenic plants

Abstract

Immunoglobulin light and heavy chains are synthesized in mammalian cells as precursors containing a signal peptide. Processing and assembling result in formation of active antibodies. Chimeric genes have been made containing the coding sequence of the barley α -amylase signal peptide which has been fused to cDNAs coding for either the mature light or the mature heavy chain of a monoclonal antibody. A plasmid was constructed linking both chimeric genes under the control of plant active promoters in an expression cassette. This DNA fragment was stably integrated into the genome of *Nicotiana tabacum* by *Agrobacterium tumefaciens* mediated gene transfer. Synthesis of light and heavy chains and assembly to antibodies was detected in transgenic tobacco tissue using specific secondary antibodies. By electron microscopic immunogold labeling, the presence of assembled antibody could be detected within the endoplasmic reticulum. Affinity chromatography indicated biological activity of the assembled immunoglobulin produced in plant cells. Unexpectedly, a significant amount of assembled antibodies was found within chloroplasts.

Introduction

Expression of light and heavy chains of monoclonal antibodies in non-lymphoid tissues has become an important area of research during the past years. In native lymphoid cells, immunoglobulin light and heavy chains are synthesized as precursor proteins. The signal peptides direct insertion of the nascent chains into the lumen of the endoplasmic reticulum (ER). Inside the ER, a heavy chain binding protein (BiP) interacting with

noncovalently associated immunoglobulin heavy chains mediates assembly of immunoglobulin light and heavy chains [4, 17, 20]. Recently a relationship of BiP to other proteins regulating transport of secretory and membrane proteins (GRP 78 and HSP 70) was reported. These proteins seem to have a general function in stress management [12, 18, 21, 26].

In *E. coli*, expression of light and heavy chains of monoclonal antibodies lacking the natural signal peptides did not result in the production of

THIS PAGE BLANK (USPTO)

functional antibodies within the transformed bacterial cells. Only *in vitro* reconstitution of bacterial extracts or purified fractions containing both chains resulted in the recovery of antigen binding activity [6, 9]. This was ascribed to the highly reducing intracellular environment inhibiting disulfide-bond formation or the lack of a special helper protein [9] or simply to low level expression [5]. Expression of a native light chain precursor gene produced a secreted immunoglobulin chain [46]. Recently, fusion of both intact or truncated light and heavy chains to bacterial signal peptides and expression of these chimeric genes in *E. coli* was successful. Protein engineering allowed the assembly of secreted functional chimeric antibodies or fragments linked by disulfide bonds [1, 35].

The immunoglobulin chains expressed by Boss *et al.* [5] in *E. coli* were also expressed in yeast cells but in this case as native precursor proteins [43]. Synthesis, processing and secretion of light and heavy chain, glycosylation of heavy chain and detection of functional antibodies in extracts from cells co-expressing both chains was demonstrated. Yet, the efficiency of assembly was low. By fusion of chimeric immunoglobulin chains to the yeast invertase signal peptide Horwitz *et al.* [22] achieved secretion of properly folded and assembled chimeric antibody and F_{ab}-fragment from yeast cells.

We have constructed chimeric genes consisting of the mature light and heavy chain genes, derived from cDNA clones of the B 1-8 antibody [6, 7] and the barley aleurone α -amylase signal peptide coding sequence [13, 14]. This signal peptide has already been used successfully to direct transport of bacteriophage T4 lysozyme from transgenic tobacco cells to the intercellular spaces [15, 23]. Thus, initiation of the secretory pathway was also expected for the chimeric immunoglobulin chains in transgenic plant tissue.

Transgenic tobacco tissue synthesizing both immunoglobulin chains was shown to produce functionally active B 1-8 antibody by using specific interaction with its antigen. Subcellular localization indicates assembly of the antibody in the endoplasmic reticulum. Synthesis and assem-

bly of a complex foreign protein to a biologically active molecule inside transgenic plant cells is possible by fusing the individual chains to a plant signal peptide.

Materials and methods

All cloning and DNA manipulations were performed according to Maniatis *et al.* [30].

Plasmids, bacterial strains, and media

Plasmids and bacterial strains are listed in Table 1. pAB λ 1-15 and pAB μ -11 were gifts from A. Radbruch; pLC 1-4, pHc 1-3, pHc 3-19 and pLGV 2385 *Bgl* \rightarrow *Hind* from M. Stieger.

Recombinations, transformations and conjugations were carried out as described elsewhere [15].

Construction of chimeric signal peptide light- and heavy-chain genes

The DNA sequence coding for the signal peptide of α -amylase from barley aleurone layer (*Clone E* [33]) was fused to the cDNA fragment coding for the mature light chain of B 1-8 antibody [8] (Fig. 1b). The signal peptide coding fragment was excised from pSR 1-1, a pUC 9 plasmid containing the *Clone E* α -amylase cDNA of Rogers and Milliman [14]. A synthetic oligonucleotide providing a 5' *Hind* III site was ligated under conservation of the *Nco* I site which overlaps with the ATG initiation codon of the signal peptide. The mature light chain coding sequence was excised from plasmid pLGV 2385 Lc. The isolated signal peptide coding sequence was fused at its 3' end to a synthetic oligonucleotide providing the 3' nucleotides of the signal peptide coding sequence and the 5' nucleotides of the light chain gene. The complete chimeric gene was ligated in the correct orientation into the selection-expression vector pAP 2034 [41] containing the T_R-1' 2' dual pro-

THIS PAGE BLANK (USPTO)

Table 1. Plasmids and bacterial strains.

Bacterial strains	Plasmids	References
<i>E. coli</i>		
DH 1		Low [29]
BMH 71-18		Yanish-Perron <i>et al.</i> [44]
GJ 23		Van Haute <i>et al.</i> [40]
<i>Agrobacterium</i>		
C58 C1		Zambryski <i>et al.</i> [45]
	pAB λ 1-15	Bothwell <i>et al.</i> [8]
	pAB μ -11	Bothwell <i>et al.</i> [7]
	pLC 1-4	Stieger [38]
	pHC 1-3	Stieger [38]
	pHC 3-19	Stieger [38]
	pLGV 2385 Lc	Stieger [38]
	pLGV 2385 Bgl \rightarrow Hind	Stieger [38]
	Clone E	Rogers and Milliman [33]
	pAP 2034	Velten and Schell [41]
	pUC 9	Vieira and Messing [42]
	pBR 322	Bollivar <i>et al.</i> [5]
	pGV 3850	Zambryski <i>et al.</i> [45]
	pSR 1-1	Düring <i>et al.</i> [15]

moter and the gene 7 polyadenylation site (derived from the T-DNA of *Agrobacterium tumefaciens*) (pSR 1-4).

An analogous construction was built with the mature heavy chain coding sequence (Fig. 1c). The signal peptide coding sequence was again prepared from plasmid pSR 1-1 but fused to another synthetic oligonucleotide at the 3' end comprising the missing 3' nucleotides of the signal peptide coding sequence and the 5' part of the gene coding for the mature heavy chain.

For integration into the plant genome a transformation vector containing both chimeric genes under control of plant active regulatory elements within a short distance was constructed. This final transformation vector contains under control of the pT_R dual promoter the NPT II gene as a selectable marker (in the 2' position) and the chimeric light chain gene (in the 1' position) and under control of the pNOS the chimeric heavy chain gene (Fig. 1a).

This plasmid was introduced into the plant transformation vector pGV 3850 [45] by homol-

ogous recombination. Recombinant *Agrobacterium* was tested for correct integration of the plasmid by Southern blotting (modified following Southern *et al.* [37], data not shown).

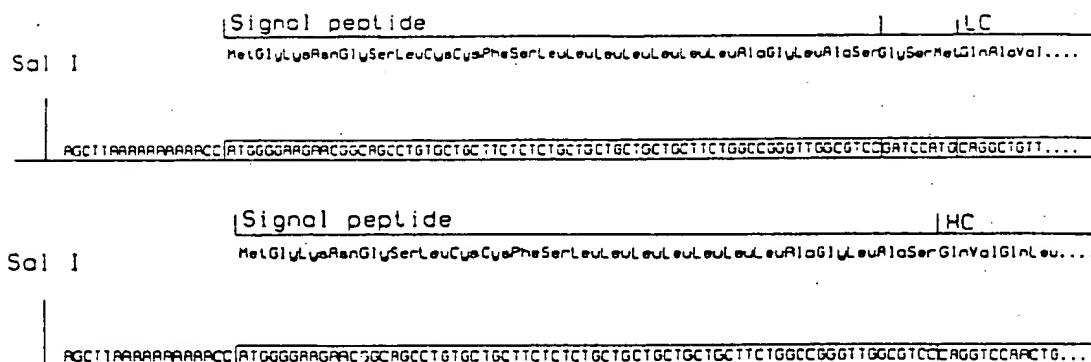
Plant transformations

Transformations of *Nicotiana tabacum* W38 plants grown in sterile culture were performed as described elsewhere [15].

Demonstration of transformation of regenerated plants

DNA analysis of transformed plants regenerated from the leaf disk infections were done according to Southern [37] and Maniatis *et al.* [30]. NPT-activity tests were performed following the protocols of Reiss *et al.* [31] and Schreier *et al.* [35].

THIS PAGE BLANK (USPTO)



Tissue printing

extract is often essential in order to circumvent cross reactions.

Protein analysis by western blotting and affinity purification of foreign protein from transformed plant tissue

B 1-8 protein, Ls 136, Ac 38, goat anti-mouse IgM and goat anti- λ antibodies as well as NP-Sepharose, L136-Sepharose and NiP-cap were gifts from A. Radbruch and M. Reth.

Protein analysis was carried out using a modi-

THIS PAGE BLANK (USPTO)

fied version of western blotting [39]. Plant tissue was homogenized in SDS-sample buffer supplemented by protease inhibitors (1 μ M leupeptin, 1 μ M pepstatin, 200 μ M PMSF, 100 μ M EDTA) and loaded on a discontinuous 15% SDS-polyacrylamide gel according to Laemmli [28]. The separated proteins were transferred to an Immobilon PVDF membrane (Millipore, Bedford, USA) by semi-dry electroblotting [27]. For immunodetection the membrane was blocked with 1% gelatin (BioRad, Richmond, USA) at room temperature to avoid nonspecific adhesion of proteins to the membrane. All antibodies were first preincubated for 10 min with wild-type tobacco protein extract to minimize nonspecific protein-protein interactions in TBS buffer (10 mM Tris-HCl pH 8.0; 150 mM NaCl) including 1% Triton X-100, 0.1% Tween 20, 0.3% gelatin and protease inhibitors. After each incubation the membrane was washed 1 \times in TBST (TBS + 0.1% Tween 20), 1 \times in TBST + 1 M NaCl and 3 \times in TBST (each 10 min).

Antibodies used for detection of individual immunoglobulin chains and assembled antibody are described in Table 2.

Table 2. Antibodies.

Monoclonal antibody	Epitope recognized on B 1-8	Reference
Ls 136	Localized only on λ -light chain	Reth [32]
Ac 38	Localized next to the active site, formed by quarternary structure of light and heavy chain	Reth [32]
Polyclonal antibody	Used for detection of:	References
Goat anti- λ	Light chain	Southern Biotechnol. Assoc.
Goat anti-IgM	Heavy chain	A. Radbruch
Bio-anti mouse IgG	(second antibody)	Sigma
Bio-anti goat IgG	(second antibody)	Sigma

Proteins were detected by a series of incubations with appropriate antibodies. Anti-IgM, Ls 136 and Ac 38 were followed by a biotinylated secondary antibody and the streptavidin-alkaline phosphatase conjugate from BRL (Bethesda, USA). In the case of biotinylated anti- λ the second antibody was omitted. Colour development became visible within a few minutes to some hours after application of NBT (Nitroblue tetrazolium) and BCIP (5-bromo-4-chloro-3-indolylphosphate) in 100 mM Tris-HCl pH 9.5, 100 mM NaCl and 5 mM $MgCl_2$.

For affinity chromatographic purification of the foreign protein from transformed plant tissue the plant material was homogenized in a mortar with a pestle under liquid nitrogen cooling. The homogenate was transferred to a Corex tube and extraction buffer (100 mM PBS) and polyvinylpolypyrrolidone (PVPP) were added. Per gram of tissue 50 mg PVPP and 1 ml 1 \times PBS supplemented with protease inhibitors (as described above) were added. As a protective agent against oxidation the solution is made 5 mM ascorbic acid. In order to solubilize all membrane encapsulated proteins 1% Triton X-100 or better 10 mM CHAPS may be added as a detergent. After 15 min extraction was finished by centrifugation of cell debris at 10000 rpm at 4 °C in a Sorvall SS 34 rotor (DuPont, Dreieich, FRG). If a detergent was used for solubilization the supernatant had to be diluted to about 0.1% Triton X-100 or 2 mM CHAPS for affinity chromatography. The supernatant was rotated on a roler at 4 °C for 30–60 min to precipitate oligosaccharides. Subsequently the precipitate was centrifuged at 4000 rpm at 4 °C. Total amount of protein was determined by a Bradford assay (BioRad, Richmond, USA) and equal amounts of total protein were used for transformed and control tissue. The protein extract was incubated with NP-Sepharose (the hapten for isolating the functional B 1-8 antibody) or Ls 136-Sepharose (for isolating the light chain) overnight at 4 °C under slow rotation in a batch assay. By centrifugation (4000 rpm, 4 °C in a Varifuge, Heraeus, Osterode, FRG) the affinity gel was precipitated and the supernatant was discarded. The gel was

THIS PAGE BLANK (USPTO)

transferred to an Eppendorf tube. All the following steps were performed at 4 °C. The gel was washed subsequently 1 × with 1 ml 1 × PBS, 1 × with 1 ml 1 × PBS, 1 M NaCl and 3 × with 1 ml 1 × PBS; all solutions were supplemented with protease inhibitors. The specifically bound antibody was eluted 3 × with 200 µl of a 10⁻⁴ M NIP-cap solution [(5-iodo-4-hydroxy-3-nitrophenyl)acetyl- α -aminocaproic acid] from the NP-Sepharose or 3 × with 200 µl of 0.1 M glycine pH 2.8 and immediately neutralized with 0.1 vol 1 M Tris-HCl pH 8.0 (for the Ls 136-Sepharose). The eluate was concentrated by ultrafiltration, desalted by washing with 1 vol water and total volume was reduced to about 10 µl. The same volume of 2 × SDS-sample buffer was added and the sample loaded onto a western gel and analyzed as described above.

Immunogold labeling

Plant tissue was prepared following the high-pressure freezing method in combination with freeze substitution and low-temperature embedding in Lowicryl HM 20 as described elsewhere [23]. Immunocytochemistry followed the same protocol but with some modifications. In some cases an additional blocking agent (3% newborn calf serum) was introduced in the preincubation step. Antibodies used for immunogold labeling were polyclonal biotinylated anti- λ , monoclonal Ls 136 and biotinylated Ls 136 for detection of the light chain and monoclonal Ac 38 for detection of the assembled B 1-8 antibody. Two systems were used for detection: either direct incubation of the first antibody with colloidal gold-adsorbed Protein A (Auroprobe EM Protein A, G 10, Janssen, Beerse, Belgium) for polyclonal antibodies and colloidal gold-adsorbed goat anti-mouse IgG second antibody (Auroprobe EM Goat anti-mouse IgG, G 10; Janssen, Beerse, Belgium) for monoclonal antibodies or colloidal gold-adsorbed streptavidin (Auroprobe EM Streptavidin, G 10, Janssen, Beerse, Belgium) for all biotinylated antibodies. In the case of Ac 38 a biotinylated goat anti-mouse IgG second antibody was used to enter the streptavidin system.

Results

Integration of chimeric immunoglobulin genes into tobacco plants

The chimeric plant signal peptide light- and heavy-chain genes (Fig. 1b and c) were cloned into a single plasmid (pSR 4-3, Fig. 1a). This vector is derived from the plant selection-expression vector pAP 2034 [41]. The chimeric light-chain gene was integrated into the single *Sal* I cloning site next to the 1' position of the pT_R dual promoter. The chimeric heavy-chain gene was set under the control of the pNOS promoter. Both promoters derive from the *Agrobacterium tumefaciens* T-DNA. The plasmid pSR 4-3 was integrated into the modified Ti-plasmid pGV 3850 [45] by homologous recombination. By leaf disk infection followed by selection on kanamycin the chimeric genes were stably integrated into the genome of *Nicotiana tabacum*. Integrity of the transferred DNA was confirmed by Southern blotting (data not shown). The functionality of the dual promoter in transgenic plant tissues was confirmed by testing for activity of the NPT II gene [31, 35] present in the 2' position of the dual promoter (data not shown).

Detection of immunoglobulin chains and assembled antibodies by affinity chromatography and immunoblotting

For detection and characterization of the monoclonal antibody B 1-8 a set of different anti-antibodies was used (Table 2). Apart from the polyclonal antibodies interacting specifically with light or heavy chains, several monoclonal anti-idiotypic antibodies are available [32]. Ls 136 recognizes the lambda light chain individually as well as in the assembled antibody. The idiotope is determined only by the light chain. Ac 38 can be used to detect the correct assembly within plant cells because the idiotope is formed by parts both of V_H and V_L domains. The idiotope is not part of the antigen binding site but located closely to it. No interaction of Ac 38 does occur with iso-

THIS PAGE BLANK (USPTO)

lated light or heavy chains. The idiotope is presumably located within the hypervariable regions. All the six hypervariable regions of B 1-8 form a cavity around the NP-binding site.

By 'tissue printing' [10] presence of light and heavy chain as well as the assembly to intact antibodies was indicated using different polyclonal and monoclonal anti-antibodies. Callus tissue derived from leaves of transformed plants was used for analysis because activity of both promoters was shown to be highest in calli (K. Düring, unpublished results). In transformed tissue the blue color of the enzymatic reaction overlaid the light-green color originating from chlorophyll partially fixed to the filters. Using

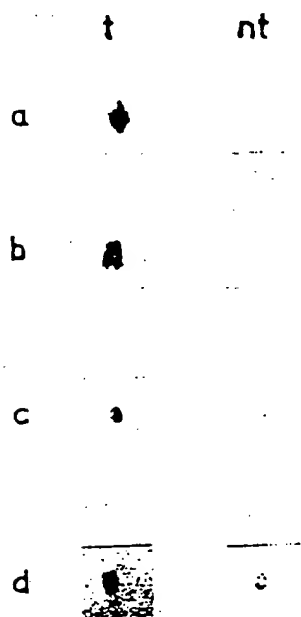


Fig. 2. 'Tissue printing' of callus tissue regenerated from transformed tobacco plants as described in 'Materials and methods'. Weak stainings in the control experiments are due to light-green color originating from chlorophyll in contrast to the dark staining of the blue NBT, BCIP-substrate precipitate in printings from transformed tissue. a. Detection of light chain with biotinylated polyclonal anti- λ antibody. b. Detection of light chain with monoclonal Ls 136 antibody. c. Detection of heavy chain with polyclonal anti-mouse IgM antibody. d. Detection of assembled antibody with monoclonal Ac 38 antibody (t: transformed tissue, nt: wild-type (non-transformed) tissue as control).

polyclonal anti- λ (Fig. 2a) and monoclonal Ls136 (Fig. 2b) the presence of light chain and using goat anti-mouse IgM (Fig. 2c) the presence of heavy chain in transgenic callus tissue was demonstrated. The specific binding of Ac 38 (Fig. 2d) to the plant proteins adsorbed to the membrane indicates assembly of both chains to structurally intact antibodies providing the correct tertiary structure of the idiotope formed by both chains.

Western blotting using crude protein extracts was successful only in the case of callus tissue. Detection by the polyclonal anti- λ antibody revealed synthesis and processing of the chimeric precursor light chain (Fig. 3a). The observed band migrated at the same position as purified B 1-8 light chain (which is the processed form). No additional band providing a higher molecular mass which could represent a precursor light chain could be detected. This suggests that processing of the foreign light-chain precursor protein in plant cells is very efficient.

Analysis of leaf tissue was performed after wounding and incubation for 5 days with 10^{-5} M 2,4-D and 10^{-5} M Kinetin. This was done in order to obtain highest possible induced expression rates for both foreign genes (K. Düring, unpublished results). Purification by affinity chromatography with the NP hapten applied to protein extracts from induced greenhouse plant leaves confirmed the presence of assembled B 1-8. Binding of the antibody produced in plant cells to its antigen also implies its biological activity. Plant material was extracted and total protein incubated with a NP-Sepharose affinity gel. The binding affinity of B 1-8 to NiP (a related hapten; heteroclitic properties of B 1-8) is higher than to NP (the original immunizing hapten) [32]. Elution of the affinity gel with NiP released bound antibodies which were then loaded onto a denaturing western gel. The B 1-8 light chain could be detected by the polyclonal anti- λ antibody. This indirect analysis proves the functionality of the assembled antibody in plant extracts because isolated light chains do not interact with NP.

Extraction of plant tissue with a pure PBS solution produced a cross-reacting band on the

THIS PAGE BLANK (USPTO)

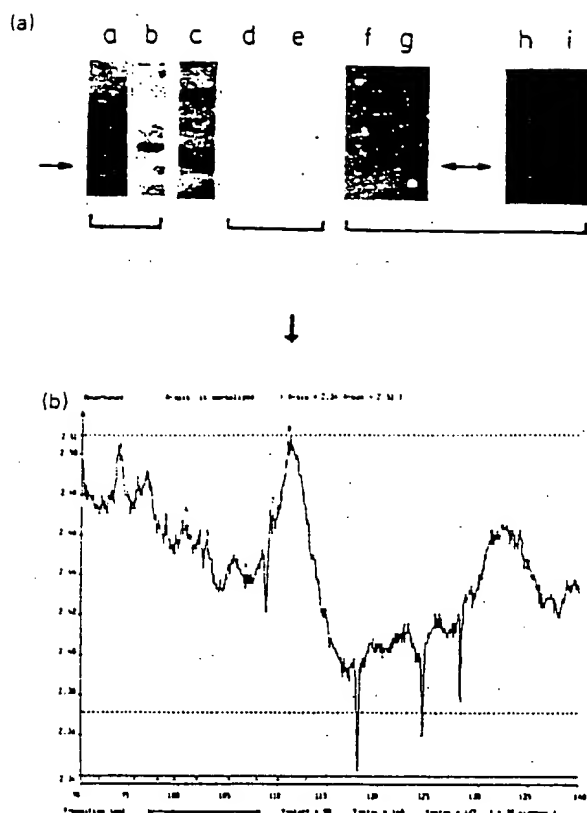


Fig. 3. a (top). Western blotting of transgenic tobacco tissue as described in 'Materials and methods', detailed description of experiments in 'Results'. Lane a: crude extract from callus regenerated from transformed leaf; lane b: crude extract from wild-type W38 callus; lane c: positive control: purified B 1-8 protein, processed light chain; lane a-c: all samples were extracted in denaturing SDS-sample buffer at 95 °C; detection of light chain with polyclonal anti- λ antibody; only processed light chain can be detected in transgenic tissue; lane d: demonstration of functionality of B 1-8 produced in induced transgenic greenhouse tobacco leaves by affinity chromatographic purification of the antibody with NP-Sepharose from plant PBS extract and subsequent denaturing western blotting; detection of light chain with polyclonal anti- λ antibody; upper band shows processed light chain; lower band: non-specific cross reaction; lane e as lane d: negative control with wild-type W38 leaves; lanes f and g: analysis of 10 mM CHAPS-PBS extract from induced greenhouse leaves by affinity chromatographic purification of the antibody and subsequent denaturing western blotting; detection of light chain with polyclonal anti- λ antibody; lane f: isolation of light chain with Ls 136-Sepharose; lane g: demonstration of functionality of B 1-8 produced in transgenic plant tissue with NP-Sepharose and detection of light chain; lane h and i as lanes f and g, respectively: negative controls with wild-type W38 leaves.

blot slightly lower than light chain which is also present in the wild-type tissue control (Fig. 3a, lane d + e). Addition of the detergent CHAPS to the extraction buffer prevented binding of this protein to the affinity gel but decreased band intensity on the blot (Fig. 3a, lane g + i). Additionally, the detection of light chain was confirmed by affinity chromatography from a CHAPS-PBS extract using the Ls 136 monoclonal antibody (Fig. 3a, lane f + h). Analysis of the blotting membrane (Fig. 3, lane g) by densitometer scanning corroborated the results obtained by eye visualization (Fig. 3b). Also in these experiments no precursor light chain could be detected. All bands detected were located at the same position as purified light chain.

Localization by electron microscopic immunogold labeling

Callus and stem tissue from transgenic tobacco plants induced by wounding and hormone application were analyzed by immunogold labeling at electron microscopic level following low-temperature preparation [23]. Ultrathin sections from callus tissue were incubated with biotinylated Ls 136 antibody and streptavidin-gold. Labeling was detected as striking accumulations of 3 to 10 gold particles in the cytoplasm (Fig. 4c) which were not found in the controls of wild-type tissue. Accumulation of several gold particles is typical of the biotin/streptavidin enhancement technique. Sections from callus and induced stem tissue were labeled with the monoclonal Ac 38 antibody and detected by colloidal gold-adsorbed anti-mouse IgG antibody. Assembled B 1-8 antibody could be detected in the endoplasmic reticulum (Fig. 4a) and surprisingly also inside chloroplasts often localized at thylakoid membranes (Fig. 4b; mostly 10–20 gold particles per chloroplast).

b (bottom). Densitometer scanning of the western blotting membrane, Fig. 3a lane g. Analysis was done on a LKB UltraScan XL laser densitometer (LKB, Bromma, Sweden); measurement without smoothing; the peak corresponding to the size of mature processed light chain is indicated by an arrow.

THIS PAGE BLANK (USPTO)

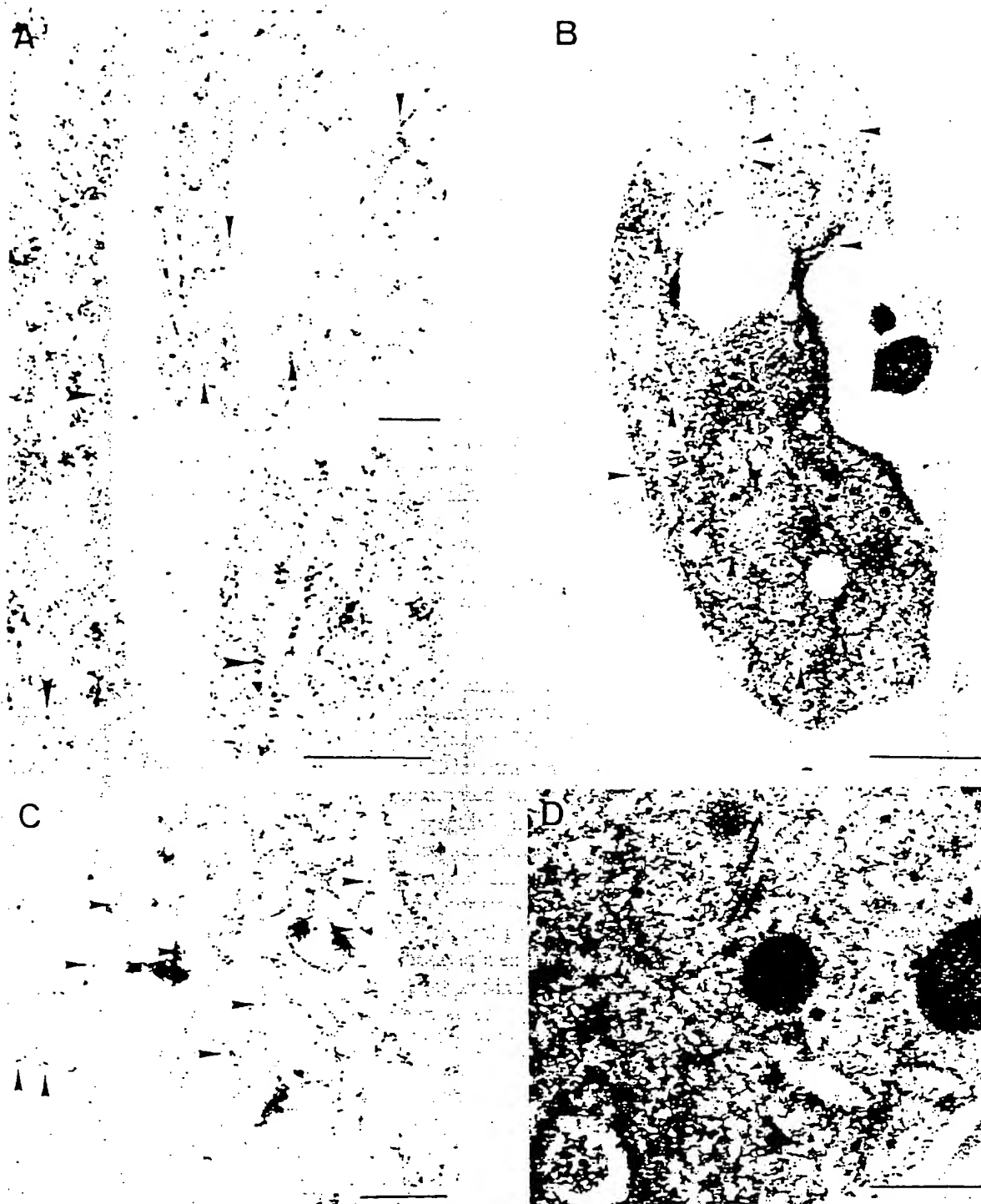


Fig. 4. Immunocytochemical localization of light chain and assembled antibody in sections of callus and induced stem tissue from transgenic tobacco plants. For details of electron microscopy see 'Materials and methods'.

A-C: transformed tissue; A: callus: labeled endoplasmic reticulum, detection of assembled antibody with monoclonal Ac 38 antibody; inset: enlarged overview on the endoplasmic reticulum section at lower magnification; B: lowest stem, induced: labeled chloroplast, detection of assembled antibody with monoclonal Ac 38 antibody; C: callus: cytoplasmic labeling, detection of light chain with monoclonal Ls 136 antibody; D: wild-type stem tissue: cytoplasm with endoplasmic reticulum, labeled with Ac 38; no unspecific labeling detected; bars = 0.5 μm.

THIS PAGE BLANK (USPTO)

Controls did not show any specific labeling. No labeling could be detected near or within cell walls or in the intercellular spaces, in the Golgi apparatus or in secretory vesicles. Fine structure localization will be continued.

Assembly of the B 1-8 antibody in plant cells inside the endoplasmic reticulum is indicated by specific labeling. Detection of light chain in the cytoplasm without any corresponding labeling for assembled antibody suggests that part of the synthesized light chain is separated from the normal secretory pathway.

Discussion

Immunoglobulins are synthesized in mammalian cells as precursors, containing specific signal peptides. After transport to the ER lumen these N-terminal extensions are cleaved off. The heavy chain is then glycosylated and light and heavy chains assemble to form intact antibodies.

For expression and assembly of the immunoglobulin subunits we constructed chimeric genes containing a plant signal peptide and the structural genes [13, 14]. A similar approach has already been taken for secretion of a chimeric T4 lysozyme in transgenic plants [15, 23]. The barley α -amylase signal peptide effected transport to the intercellular space. Using the same signal peptide we expected to find the immunoglobulin chains within the ER lumen. Comparable experiments in *E. coli* and yeast have shown that assembly of an active antibody is possible in both organisms only in the presence of signal peptides in the foreign precursors [1, 24, 36, 43]. Our data provide evidence that light and heavy chain are transported to the ER and that the subunits assemble within its lumen.

Detection of the synthesis and assembly of the monoclonal B 1-8 antibody in transgenic tobacco cells was performed by immunological methods. Tissue printing and western blotting provided data showing synthesis of both immunoglobulin chains as well as processing of the chimeric precursor light chain. The heavy chain could not be detected on western blots because of the lack of

a specific and highly sensitive anti-antibody. Assembly of the antibody could be shown by tissue printing using an anti-idiotypic antibody. Binding of B 1-8 produced by plant cells to its Sepharose-bound hapten (NP) indicates functional activity. The bound antibody can be released by a related hapten (NiP) because B 1-8 shows heteroclitic properties. Detection of the light chain in this eluate demonstrates the presence of a functional antibody within plant cells. Individual light chains do not interact with the NP hapten [32]. Assembly of light and heavy chains in the ER is indicated by immunogold labeling using the Ac 38 anti-idiotypic antibody. Glycosylation of the heavy chain has not been investigated because of insufficient sensitivity.

Assembly of a stable active antibody involves formation of intramolecular disulfide bonds. This should be greatly facilitated in the endoplasmic reticulum because of the presence of disulfide isomerase. As described for a eucaryotic protein expressed in a procaryotic cell, the bacterial periplasmic space corresponding to the eucaryotic endoplasmic reticulum in the secretory pathway provides an environment facilitating formation of disulfide bridges since both compartments are involved in the first step of protein export. Spatial concentration within the ER may promote assembly to an active antibody. In addition, the important role of chaperonins such as BiP in protein folding and assembly is currently to be elucidated [6, 12, 19, 20, 26, 34]. Most probably similar plant chaperonins can be expected also in the ER but not in the cytoplasm. All these factors may be important in producing multimeric proteins, as antibodies in transgenic plants. Nevertheless, in mouse cells the assembly of B 1-8 antibody lacking any signal peptide was obtained cytoplasmically although with very low efficiency [3].

Immunogold labeling at electron microscopic level provided additional data for the localization in the transgenic tissue. Binding of monoclonal Ls 36 antibody in the cytoplasm has to be due to the presence of individual light chains since no corresponding labeling could be found for an assembled antibody. This labeling may represent

THIS PAGE BLANK (USPTO)

a part of the light chain aborted from the secretory pathway. Garcia *et al.* [16] described an analogous phenomenon for the precore protein of hepatitis B virus. About 70–80% of the protein were found in an *in vitro* experiment in the cytosolic fraction although the signal peptide had been cleaved off implying a release of the processed protein from the endoplasmic reticulum. Immunoglobulin light chain may be produced in excess in relation to the heavy chain in transformed tissue because of higher promoter activity of the dual promoter directing the chimeric light-chain gene. Therefore part of the light-chain molecules always may remain unbound.

Localization of assembled antibodies in chloroplasts was detected unexpectedly. Significant amounts of gold particles were found with the monoclonal Ac 38 antibody inside these organelles. Control experiments with wild-type tissue did not show unspecific labeling. No convincing explanation can be given at the actual knowledge. No features of these engineered proteins were expected to direct an import of assembled antibody or individual immunoglobulin chains into chloroplasts. Therefore we cannot decide whether an import occurred or if other reasons cause this specific labeling. Immunogold labeling of transgenic tobacco tissue secreting a chimeric T4 lysozyme fused to the same α -amylase signal peptide revealed an unexpected localization of lysozyme in proplastids [15]. Elucidation of possible correlations requires further analysis.

No other labelings characterizing the secretory pathway or showing a final localization of the assembled antibody or its individual chains were found yet. Further fine structural analysis at electron microscopic level is under work.

In conclusion, the experiments described here provide evidence for the synthesis, insertion into the endoplasmic reticulum and assembly of immunoglobulin light and heavy chains as well as for the biological activity of the produced monoclonal antibody. Fusion of the immunoglobulin chains to a plant signal peptide may be the essential factor for production in transgenic plants when reviewing the actual state of knowledge concerning biosynthesis of antibodies in lymphoid

cells. Expression of monoclonal antibodies in plants might be of importance in basic as well as in applied plant research.

In a paper just published by Hiatt *et al.* [22] a very similar approach for expression and assembly of a monoclonal antibody in transgenic plant tissue was described. In this case the natural mammalian signal peptides of light and heavy chains were used instead of a plant originating signal peptide. Analogous constructions without any signal peptide were also tested but only very low levels of single chains and no assembly could be detected in transgenic plant tissue. It was reported that expression of both precursor coding cDNAs in one plant resulted in formation of high levels of assembled antibodies. Data were obtained by ELISA and non-denaturing western blotting but not by affinity chromatography. The two chimeric genes were first introduced individually into different tobacco plants and then combined by sexual crossing. Parental plants expressing only one chain showed low levels of foreign protein. F1 plants expressing both chains revealed high levels of assembled antibodies. As RNA levels in all these plants were not significantly different light- and heavy-chain proteins have to be stabilized by aggregation.

Both strategies to assemble active monoclonal antibodies in transgenic plants were successful. Therefore we can conclude that the most important factor is the presence of a signal peptide fused to the mature light- and heavy-chain proteins and not its origin. In addition to the work described by Hiatt *et al.* we could demonstrate purification of the functional assembled antibody by affinity chromatography and localization by electron microscopic immunogold labeling. It was demonstrated that assembly indeed occurs in the endoplasmic reticulum also in transgenic plants. Both approaches may be useful also for assembly of other multimeric foreign proteins in plant tissue.

Acknowledgements

We thank Inge Czaja, Ulrike Haus, Monika Hermanns and Heidi Schlichtherle for technical

THIS PAGE BLANK (USPTO)

help and Dr. Robert Masterson for many valuable discussions. Our special thanks are due to Prof. Dr. Hans Moor for providing the facilities of using a Balzers HPM 010 high-pressure freezing machine. This work was supported by a predoctoral grant to K.D. by the Fonds der Chemischen Industrie Deutschlands.

References

1. Better M, Chang CP, Robinson RR, Horwitz AH: *Escherichia coli* secretion of an active chimeric antibody fragment. *Science* 240: 1041-1043 (1988).
2. Bevan M, Barnes WM, Chilton MD: Structure and transcription of the nopaline synthase gene region of T-DNA. *Nucl Acids Res* 11: 369-385 (1983).
3. Biocca S, Neuberger MS, Cattaneo A: Expression and targeting of intracellular antibodies in mammalian cells. *EMBO J* 9: 101-108 (1990).
4. Bole DG, Hendershot L, Kearney JF: Posttranslational association of immunoglobulin heavy chain binding protein with nascent heavy chains in non-secreting and secreting hybridomas. *J Cell Biol* 102: 1558-1566 (1986).
5. Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Heyneker HL, Boyer HW: Construction and characterization of new cloning vehicles II: A multipurpose cloning system. *Gene* 2: 95-113 (1977).
6. Boss MA, Kenten JH, Wood CR, Emtage SJ: Assembly of functional antibodies from immunoglobulin heavy and light chains synthesized in *E. coli*. *Nucl Acids Res* 12: 3791-3806 (1984).
7. Bothwell ALM, Paskind M, Reth M, Imanishi-Kari T, Rajewski K, Baltimore D: Heavy chain variable region contribution to the NP^b family of antibodies: somatic mutation evident in a γ 2a variable region. *Cell* 23: 625-637 (1981).
8. Bothwell ALM, Paskind M, Reth M, Imanishi-Kari T, Rajewski K, Baltimore D: Somatic variants of murine immunoglobulin light chains. *Nature* 298: 380-382 (1982).
9. Cabilly S, Riggs AD, Pande H, Shively JE, Holmes WE, Rey M, Perry LJ, Wetzel R, Heyneker HL: Generation of antibody activity from immunoglobulin polypeptide chains produced in *E. coli*. *Proc Natl Acad Sci USA* 81: 3273-3277 (1984).
10. Cassab GI, Varner J: Immunocytochemical localization of extensin in developing soybean seed coats by immunogold-silver staining and by tissue printing on nitrocellulose paper. *J Cell Biol* 105: 2581-2588 (1987).
11. Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM: Nopaline synthase: transcript mapping and DNA sequence. *J Mol Appl Genet* 1: 561-573 (1982).
12. Dorner AJ, Bole DG, Kaufman RJ: The relationship of N-linked glycosylation and heavy-chain binding protein association with the secretion of glycoproteins. *J Cell Biol* 105: 2665-2674 (1987).
13. Düring K: Wundinduzierbare Expression und Sekretion von T4 Lysozym und monoklonalen Antikörpern in *Nicotiana tabacum*. Ph.D. thesis, Universität Köln, FRG (1988).
14. Düring K, Hippe S: Synthesis, assembly and targeting of foreign chimeric proteins in transgenic *Nicotiana tabacum* cells. Abstract 'Herbsttagung der Gesellschaft für Biologische Chemie 1989 Osnabrück'. *Biol Chem Hoppe-Seyler* 370: 888 (1989).
15. Düring K, Hippe S, Kreuzaler F, Schell J: Transport of a chimeric bacteriophage T4 lysozyme to the intercellular spaces in transgenic tobacco plants. submitted.
16. Garcia PD, Ou JH, Rutter WJ, Walter P: Targeting of the hepatitis B virus precore protein to the endoplasmic reticulum membrane: after signal peptide cleavage translocation can be aborted and the product released into the cytoplasm. *J Cell Biol* 106: 1093-1104 (1988).
17. Haas IG, Wabl M: Immunoglobulin heavy chain binding protein. *Nature* 306: 387-389 (1983).
18. Haas IG, Meo T: cDNA cloning of the immunoglobulin heavy chain binding protein. *Proc Natl Acad Sci USA* 85: 2250-2254 (1988).
19. Hemmingsen SM, Woolford C, van der Vies SM, Tilly K, Dennis DT, Georgopoulos CP, Hendrix RW, Ellis RJ: Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* 333: 330-334 (1988).
20. Hendershot L, Bole D, Köhler G, Kearney JF: Assembly and secretion of heavy chains that do not associate post-translationally with immunoglobulin heavy chain-binding protein. *J Cell Biol* 104: 761-767 (1987).
21. Hendershot LM, Ting J, Lee AS: Identity of the immunoglobulin heavy-chain-binding protein with the 78,000-Dalton glucose-regulated protein and the role of post-translational modifications in its binding function. *Mol Cell Biol* 8: 4250-4256 (1988).
22. Hiatt A, Cafferkey R, Bowditch K: Production of antibodies in transgenic plants. *Nature* 342: 76-78 (1989).
23. Hippe S, Düring K, Kreuzaler F: In situ localization of a foreign protein in transgenic plants by immunoelectron microscopy following high pressure freezing, freeze substitution and low temperature embedding. *Eur J Cell Biol* 50: 230-234 (1989).
24. Horwitz AH, Chang CP, Better M, Hellstrom KE, Robinson RR: Secretion of functional antibody and Fab fragment from yeast cells. *Proc Natl Acad Sci USA* 85: 8678-8682 (1988).
25. Hsiung HM, Mayne NG, Becker GW: High-level expression, efficient secretion and folding of human growth hormone in *Escherichia coli*. *Biotechnology* 4: 991-995 (1986).
26. Kassenbrock CK, Garcia PD, Walter P, Kelly RB:

THIS PAGE BLANK (USPTO)

- Heavy-chain binding protein recognizes aberrant polypeptides translocated in vitro. *Nature* 333: 90-93 (1988).
27. Kyhse-Andersen J: Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *J Biochem Biophys Meth* 10: 203-209 (1984).
 28. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-685 (1970).
 29. Low B: Formation of merodiploids in matings with a class of Rec⁻ recipient strains of *E. coli* K12. *Proc Natl Acad Sci USA* 60: 160-167 (1968).
 30. Maniatis F, Fritsch EF, Sambrook J: *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1982).
 31. Reiss B, Sprengel R, Will H, Schaller H: A new and sensitive method for qualitative and quantitative assay of neomycinphosphotransferase in crude cell extracts. *Gene* 30: 211-218 (1984).
 32. Reth M: Charakterisierung individueller Antikörper und Idiotope des NP^o Idiotope. Ph.D. thesis, Universität Köln, FRG (1981).
 33. Rogers JC, Milliman C: Isolation and sequence analysis of a barley α -amylase cDNA clone. *J Biol Chem* 258: 8169-8174 (1983).
 34. Roy H: Rubisco assembly: a model system for studying the mechanism of chaperonin action. *Plant Cell* 1: 1035-1042 (1989).
 35. Schreier PH, Seftor EA, Schell J, Bohnert HJ: The use of nuclear encoded sequences to direct the light regulated synthesis and transport of a foreign protein into plant chloroplasts. *EMBO J* 4: 25-32 (1985).
 36. Skerra A, Plückthun A: Assembly of a functional immunoglobulin F₁ fragment in *Escherichia coli*. *Science* 240: 1038-1041 (1988).
 37. Southern EM: Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98: 503-517 (1975).
 38. Stieger M: Versuche zur Integration und Expression chimärer Immunglobuline in Pflanzen. Ph.D. thesis, Universität Köln, FRG (1987).
 39. Towbin H, Staehelin T, Gordon J: Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci USA* 76: 4350-4354 (1979).
 40. Van Haute E, Joos H, Maes M, Warren G, Van Montagu M, Schell J: Intergenic transfer and exchange recombination of restriction fragments cloned in pBR 322: a novel strategy for the reversed genetics of Ti-plasmids of *A. tumefaciens*. *EMBO J* 2: 411-418 (1983).
 41. Velten J, Schell J: Selection-expression plasmid vectors for use in genetic transformation of higher plants. *Nucl Acids Res* 13: 6981-6998 (1986).
 42. Vieira J, Messing J: The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19: 259-268 (1982).
 43. Wood CR, Boss MA, Kenten JH, Calvert JE, Roberts NA, Emtage JS: The synthesis and in vivo assembly of functional antibodies in yeast. *Nature* 314: 446-449 (1985).
 44. Yanish-Perron C, Vieira J, Messing J: Improved M13 cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC 19 vectors. *Gene* 33: 103-119 (1985).
 45. Zambryski P, Joos H, Genetello C, Van Montagu M, Schell J: Ti-plasmid vector for introduction of DNA into plant cells without altering their normal regeneration capacity. *EMBO J* 2: 2143-2150 (1983).
 46. Zemel-Dreassen O, Zamir A: Secretion and processing of an immunoglobulin light chain in *Escherichia coli*. *Gene* 27: 315-322 (1984).

THIS PAGE BLANK (USPTO)